



Universidade de Brasília

**Instituto de Ciências Exatas
Departamento de Ciência da Computação**

Fotografias padronizadas utilizando dispositivo móvel

Matheus Rosendo Pedreira

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Engenharia da Computação

Orientador
Prof. Dr. Alexandre Zaghetto

Brasília
2016

Universidade de Brasília — UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Bacharelado em Engenharia da Computação

Coordenador: Prof. Dr. Ricardo Pezzuol Jacobi

Banca examinadora composta por:

Prof. Dr. Alexandre Zaghetto (Orientador) — CIC/UnB

Prof. Dr. Flávio de Barros Vidal — CIC/UnB

Prof. Dr. Marcus Vinícius Lamar — CIC/UnB

CIP — Catalogação Internacional na Publicação

Pedreira, Matheus Rosendo.

Fotografias padronizadas utilizando dispositivo móvel / Matheus Rosendo Pedreira. Brasília : UnB, 2016.

101 p. : il. ; 29,5 cm.

Monografia (Graduação) — Universidade de Brasília, Brasília, 2016.

1. celular, 2. padrão de fotografia, 3. reconhecimento facial,
4. reconhecimento de fundo, 5. iPhone, 6. tempo real

CDU 004.4

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro — Asa Norte
CEP 70910-900
Brasília-DF — Brasil

Dedicatória

Dedico este trabalho à minha família, meus amigos, meus sócios, à empresa júnior de engenharia da computação (Struct), à antiga gestão do CA, à LoopCE e a todos que estiveram comigo nessa jornada que foi estudar na UnB. Sem qualquer um de vocês o meu caminho na UnB teria sido mais difícil e com certeza todos acrescentaram muito ao meu crescimento acadêmico.

Agradecimentos

Agradeço ao dr. Alexandre Zaghetto por me auxiliar a criar este projeto, por toda sua orientação e sua vontade de me ajudar.

Aos meus amigos João Vitor Gontijo, Felipe Milhomem, Tiago Pigatto, André Bauer, Mateus Alves, Matheus Antunes, Rafael Lages, Vanessa Diniz e Arthur Teixeira por cederem suas imagens para que eu pudesse realizar a pesquisa que foi tão importante para o meu trabalho.

Agradeço também a todos os professores do departamento de Ciência da Computação e de Engenharia Elétrica pois sem o conhecimento adquirido, durante as diversas aulas que atendi, esse trabalho teria sido muito mais difícil de ser concluído.

Resumo

Esse trabalho foi desenvolvido pensando no crescente uso dos dispositivos móveis para se adaptar ao momento atual da sociedade, na qual a necessidade de mão de obra humana vem se tornando menor, principalmente para tarefas simples, como tirar fotografias.

O objetivo do trabalho é desenvolver um sistema onde a pessoa que o utilize seja capaz de tirar uma fotografia sozinha, com essa independência, um órgão que exija certo nível de qualidade nas fotografias tem garantida uma maior segurança em seu sistema, ao eliminar o fator humano. Para isso, foram desenvolvidos diversos sistemas para a classificação de características próprias de padrões de fotografia, quando se tratando de documentos oficiais. Tais sistemas foram unidos em uma única solução, onde o usuário é capaz de tirar uma fotografia de boa qualidade sem auxílio externo, necessitando apenas de um sistema que realiza as verificações em tempo real.

O sistema foi, em seguida, analisado utilizando métricas individuais para cada componente do sistema e como um todo através de uma pesquisa com o fim de avaliar as vantagens proporcionadas pelo mesmo quando utilizado por indivíduos sem habilidades técnicas realizando as fotografias citadas.

Palavras-chave: celular, padrão de fotografia, reconhecimento facial, reconhecimento de fundo, iPhone, tempo real

Abstract

This work was developed thinking about the growing usage of mobile devices adapting to our current moment in society, in which the need for human labor is decreasing, even more for simple tasks such as taking pictures.

The goal of this project is to develop a system where the person that is using it is capable of taking its own photograph with such standardization that an organization that requires an amount of quality from photos doesn't require the work of people specialized in analyzing the photos that were sent to them. This would then give more security and agility to their systems. To achieve that goal there were developed many systems focused on classifying characteristics used by photograph standards for official documents. These systems were then merged into one solution where the user is capable of taking a photograph with good quality and without the help of any person, only the help of a system that does the verification in real time.

The system was analyzed using individual metrics for each component for the system and then as a whole through a research intended to evaluate the advantages created from the system when compared to people without the proper knowledge to take the pictures without help.

Keywords: mobile, photograph standard , facial recognition, background recognition, iPhone, real time

Sumário

1	Introdução	1
1.1	Motivação e Justificativa	1
1.2	Objetivos	2
1.3	Organização	2
2	Revisão teórica	3
2.1	Fundamentação Teórica	3
2.1.1	Imagem digital	3
2.1.2	Codificação de imagem	4
2.1.3	Espaços de cores	4
2.1.4	Vídeo	5
2.1.5	Formato imagem/frame	5
2.1.6	Detecção rápida de objetos	6
2.1.7	Média e Variância em imagem	7
2.1.8	iPhone	8
2.2	Padrões de fotografia para documentos oficiais	8
2.2.1	Posicionamento da face	8
2.2.2	Iluminação e fundo	8
2.2.3	Formato da imagem	9
2.2.4	Informações adicionais	9
2.3	Trabalhos correlatos	10
3	Metodologia	12
3.1	Dimensões	12
3.2	Obtenção dos frames	12
3.3	Análise dos frames	12
3.4	Detecção de faces	14
3.5	Detecção de olhos	14
3.5.1	Com face	14
3.5.2	Sem face	14
3.6	Altura da face	15
3.7	Distância entre olhos	16
3.8	Distância dos olhos ao topo	16
3.9	Alinhamento dos olhos	17
3.10	Face centralizada	18
3.11	Detecção de fundo	18

3.11.1	Fundo claro	19
3.12	Ruído do fundo	20
3.13	Salvando a imagem	21
3.14	Análise dos resultados	21
4	Resultados	23
4.1	Detecção de faces	23
4.2	Detecção de olhos	24
4.2.1	Com face	24
4.2.2	Sem face	24
4.3	Altura da face	26
4.4	Distância entre olhos	26
4.5	Distância dos olhos ao topo	27
4.6	Alinhamento dos olhos	27
4.7	Face centralizada	28
4.8	Detecção de fundo	28
4.8.1	Face encontrada	28
4.8.2	Face estimada pelos olhos	29
4.9	Fundo claro	29
4.10	Ruído do fundo	30
4.11	Sistema completo	30
4.11.1	Todas as verificações desativadas x Todas as verificações ativadas	32
4.11.2	Fotos com sistema de verificação do fundo desligado	34
4.11.3	Fotos com sistema de verificação da altura da face desligado	35
4.11.4	Fotos com sistema de verificação o alinhamento dos olhos desligado	36
5	Conclusão	38
	Referências	39

Lista de Figuras

2.1	Imagens com pixel com apenas uma informação e com mais de uma informação	3
2.2	JPEG comprimido com diferentes taxas	4
2.3	Representação das cores no modelo RGB. Imagem retirada do artigo "RGB Color Spaces"(Microsoft) [13]	5
2.4	Representação das cores no modelo HSV. Imagem retirada do artigo "HSV Color Space"(Rohan Library) [12]	6
2.5	Exemplos de imagens integrais. Imagem retirada da documentação do OpenCV[18]	7
2.6	Comparação de posicionamento das faces. Imagens retiradas do guia "ICAO Guide for MRTD Photo Guidelines"[10].	9
2.7	Comparação do fundo das imagens. Imagens retiradas do guia "ICAO Guide for MRTD Photo Guidelines"[10].	9
2.8	Comparação de acessórios nas fotos. Imagens retiradas do guia "ICAO Guide for MRTD Photo Guidelines"[10].	10
3.1	Fluxograma proposto do sistema	13
3.2	Proporções da face. Figura adaptada do artigo "The proportion of the head and face"[8].	15
3.3	Distância dos olhos ao topo. Imagem adaptada de "Photos d'identité [24]."	16
3.4	O retângulo azul demonstra o alinhamento	17
3.5	Deteção de fundo. Foto exemplo de documento retirada do estúdio de fotos Delio[19]	18
3.6	Imagem base utilizada na pesquisa. Adaptada de [7].	22
4.1	Deteção de faces utilizando o classificador Haar em cascata.	23
4.2	Deteção de olhos utilizando classificador Haar em cascata na região da face encontrada.	24
4.3	Deteção olhos utilizando classificador Haar em cascata.	25
4.4	Estimação da face através dos olhos encontrados.	25
4.5	Distância entre o centro dos olhos.	26
4.6	Comparação de alinhamento dos olhos.	27
4.7	Demonstração de como o fundo é encontrado com a face encontrada.	28
4.8	Demonstração de como o fundo é encontrado com a face estimada.	29
4.9	Exemplos de imagens tiradas com o fundo colorido.	29
4.10	Imagens do sistema funcionando.	31
4.11	Imagem de referência. Adaptada de [7]	31

4.12 Imagem 1 com tudo desligado.	32
4.13 Imagem 2 com tudo desligado.	33
4.14 Imagem 1 com tudo funcionando.	33
4.15 Imagem 2 com tudo funcionando.	34
4.16 Imagem 1 com fundo desligado.	34
4.17 Imagem 2 com fundo desligado.	35
4.18 Imagem 1 com tamanho da face desligado.	35
4.19 Imagem 2 com tamanho da face desligado.	36
4.20 Imagem 1 com alinhamento desligado.	36
4.21 Imagem 2 com alinhamento desligado.	37

Lista de Tabelas

2.1	Resolução referente aos megapixels. Dados retirados do artigo "What is a megapixel?" (Digicam Help) [9]	6
2.2	Percepção de qualidade por usuários experientes	11
4.1	Razão da face em relação a imagem.	26
4.2	Distância dos olhos ao topo.	27
4.3	Tabela demonstrando a centralização da faces.	28
4.4	Verificação de claridade do fundo.	30
4.5	Verificação de ruído do fundo.	30

Lista de Símbolos

: Razão de proporção entre largura : altura. 9, 12, 15, 16

% Percentual. 4, 11, 14, 26, 28–30, 32, 34–36

Capítulo 1

Introdução

1.1 Motivação e Justificativa

Para a obtenção de determinados documentos, retratos faciais são comumente solicitados, pois são um mecanismo importante para que órgãos fiscalizadores e organizadores de concursos, por exemplo, consigam facilmente identificar as pessoas realizando uma comparação visual. Por tal motivo, o padrão ICAO 9303 foi criado e adotado parcialmente por órgãos como a embaixada dos Estados Unidos[17] e a comissão da União Europeia[4]. Esses países possuem padrões com diversas similaridades entre si, que podem ser facilmente notados.

Um fato importante é que, com o avanço da globalização, o acesso à tecnologia tem aumentado drasticamente a cada ano. Segundo o livro "Communities Dominate Brands" (AHOONEN e MOORE, 2013)[5] atualmente existem 5.2 bilhões de telefones celulares em uso no mundo; 83% deles possuem câmeras e 90% dos usuários apenas tiraram fotos com o aparelho. Devido essa grande porcentagem, foi criada uma solução de alta utilidade pública para facilitar a captura de fotografias padronizadas com aparelhos celulares para emissão de documentos.

Após verificar a maioria das aplicações móveis existentes na área como: Passport ID Photo Maker Studio (Android)[14] e ID PhotoPrint: US passport photos, delivered (iOS)[21]; Foi notado que é dependente do usuário compreender como tirar a foto, ler e interpretá-la de maneira subjetiva para concluir se a mesma está correta para, então, decidir por capturar uma nova foto ou não. Essa dependência direta do conhecimento do usuário e da capacidade de análise deste em confirmar se a foto é de qualidade, ou não, pode gerar inúmeras falhas de classificação. Entre essas falhas estão problemas como: Face não centralizada; Mais de uma face na imagem; Fundo de cor errada; Olhos fechados ou não visíveis; Imagem não focada; Tamanho da face menor ou maior do que o esperado.

Portanto, visando desenvolver uma solução para os órgãos que farão uso da fotografia com utilização de técnicas de reconhecimento facial e de análise de padrões de imagem, além de criar uma experiência melhor e mais eficiente ao usuário, que esse trabalho foi desenvolvido.

1.2 Objetivos

O objetivo do trabalho é criar uma aplicação móvel capaz de facilitar a criação de retratos faciais, que sigam os pontos mais importantes dos padrões de retratos para documentos. Para isso, é preciso otimizar a identificação de faces e análise das características necessárias para que o sistema seja capaz de funcionar de maneira rápida em um hardware de capacidades limitadas. Com o sistema desenvolvido é necessário então verificar as vantagens e desvantagens do sistema proposto sobre o que existe hoje no mercado.

1.3 Organização

O restante desse trabalho irá discutir como o sistema foi implementando, bem como a teoria necessária para entender o sistema e os experimentos realizados. O capítulo 2 têm como finalidade explicar os conceitos que foram utilizados para a realização do desenvolvimento, bem como ferramentas utilizadas e requisitos para o funcionamento. No capítulo 3 será explicado como foi realizada a implementação do software e todas suas etapas. O capítulo 4 contém as informações de resultados obtidos. Finalmente, o capítulo 5 resume o que foi analisado ao longo do trabalho e conclui sobre os resultados a fim de imaginar possíveis continuidades de pesquisa na área.

Capítulo 2

Revisão teórica

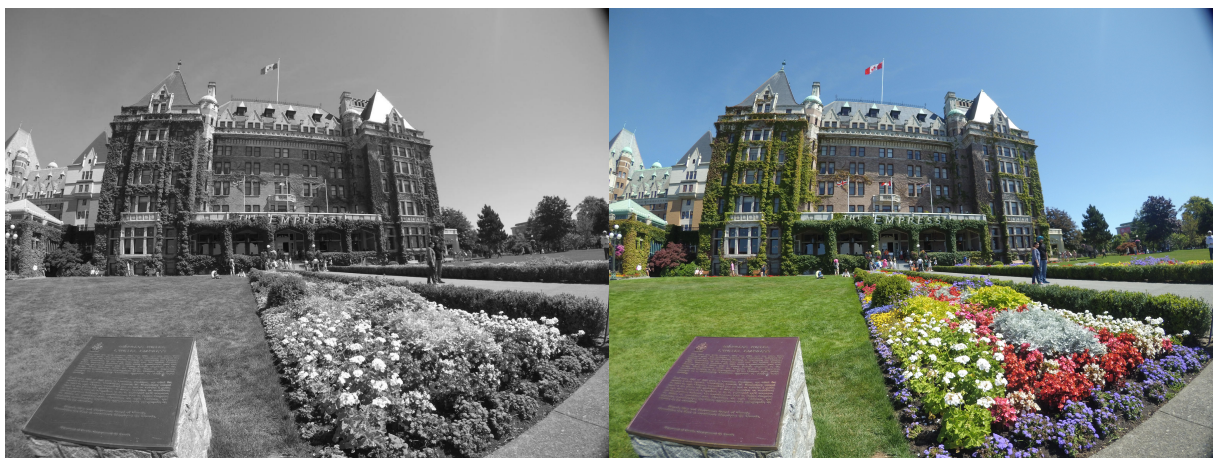
Neste capítulo serão identificados os trabalhos que foram utilizados como referência, bem como termos técnicos necessários para o entendimento deste documento e requisitos necessários utilizados no trabalho.

2.1 Fundamentação Teórica

2.1.1 Imagem digital

Como visto no livro Digital Image Processing[6], uma imagem digital é uma representação matricial de uma cena onde cada ponto é uma aproximação do ponto equivalente sendo representado. Essa matriz pode ser definida por uma função $f(x,y)$ onde o x representa a coluna e o y a linha na matriz. Cada ponto dessa matriz é chamado de pixel.

Cada pixel da imagem pode ser representado com apenas um valor, o que é comumente utilizado em processamento de imagem para analisar imagens em escala de cinza, ou pode ser representado com mais de um valor, que é normalmente utilizado para representar imagens coloridas.



(a) Imagem escala de cinza

(b) Imagem colorida

Figura 2.1: Imagens com pixel com apenas uma informação e com mais de uma informação

2.1.2 Codificação de imagem

Uma imagem pode ser salva em um arquivo sem compressão ou com compressão e essa compressão da imagem pode se dar de uma forma com perdas ou sem perdas. Uma compressão sem perdas se dá por um algoritmo que procura maneiras melhores de representar a informação de uma imagem, de tal forma que a imagem continue com toda informação inicial e reduza seu espaço utilizado em disco. Já uma compressão com perdas aceita que algumas informações sejam perdidas, assim prejudicando a qualidade da imagem para que o tamanho da imagem seja mais reduzido.

Tendo em vista que uma imagem é uma representação matricial de pixels com informações de cor, um algoritmo pode otimizar uma imagem procurando padrões nos pixels e representando eles com menos bits. Já um algoritmo com perdas poderia diminuir a quantidade de bits necessários para representar uma cor diminuindo, assim, o espectro de cores e ao mesmo tempo o tamanho da imagem.

Os dois formatos de imagem comprimidos mais utilizados são o PNG e o JPEG, sendo o primeiro sem perdas e o segundo com perdas. O primeiro utiliza da busca de padrões para otimizar o tamanho do arquivo enquanto o segundo utiliza um algoritmo, que analisa a imagem e descarta informações que ele considera que o olho humano não é capaz de perceber. No entanto esse valor de compressão do JPEG é ajustável, assim existindo vários possíveis resultados para como a compressão é realizada.

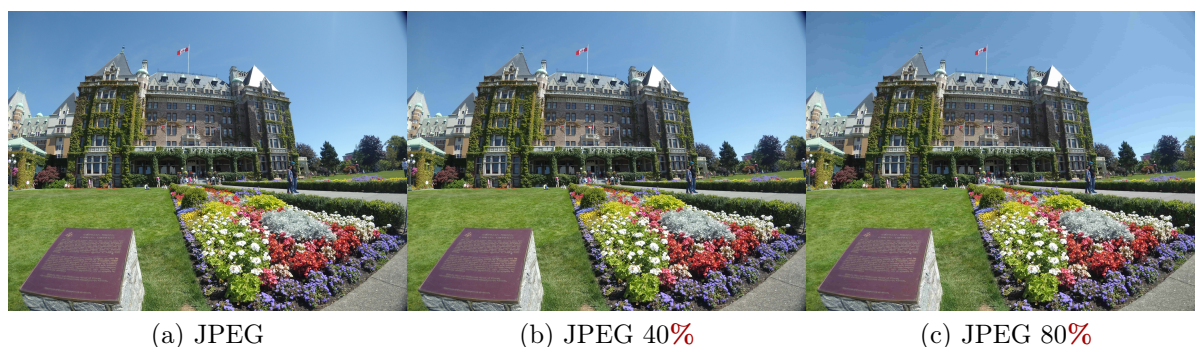


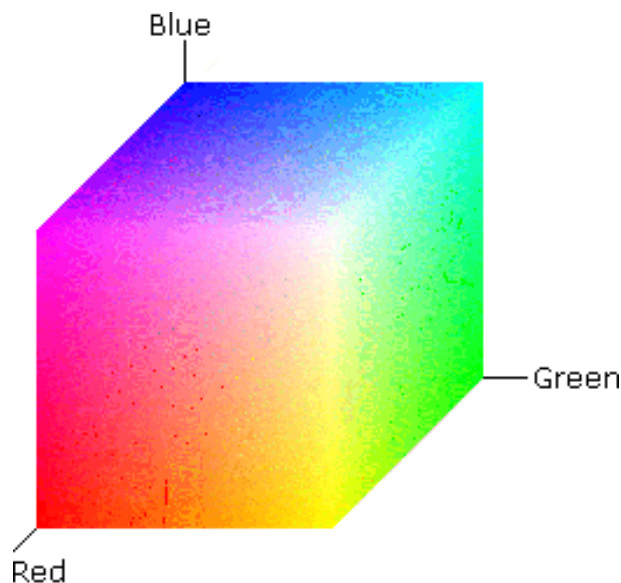
Figura 2.2: JPEG comprimido com diferentes taxas

2.1.3 Espaços de cores

Para realizar a representação de uma cor em um pixel existem diferentes modelos que podem ser utilizados. Entre esses espaços de cor temos o RGB, o HSV, entre outros. Para os fins deste trabalho os mais interessantes são o RGB e o HSV.

Como apresentado no livro Digital Image Processing[6], o modelo RGB utiliza a ideia de que qualquer cor pode ser obtida a partir de uma mistura das cores vermelho (Red), verde (Green) e azul (Blue). Portanto, esse modelo considera, que para definir a cor de um pixel é necessário que cada pixel tenha a informação de todos esses canais e considera, que o valor máximo do canal seria como se a intensidade dessa cor fosse máxima e o valor mínimo seria como se não houvesse a cor presente na mistura. Através dessa mistura é possível obter para uma representação utilizando 24 bits, 8 bits para cada canal, um total

de $(2^8)^3 = 16.777.216$ cores. Assim a cor preta pode ser representada como $R=0$, $G=0$ e $B=0$ e o branco como $R=255$, $G=255$ e $B=255$.



(a) Modelo RGB

Figura 2.3: Representação das cores no modelo RGB. Imagem retirada do artigo "RGB Color Spaces" (Microsoft) [13]

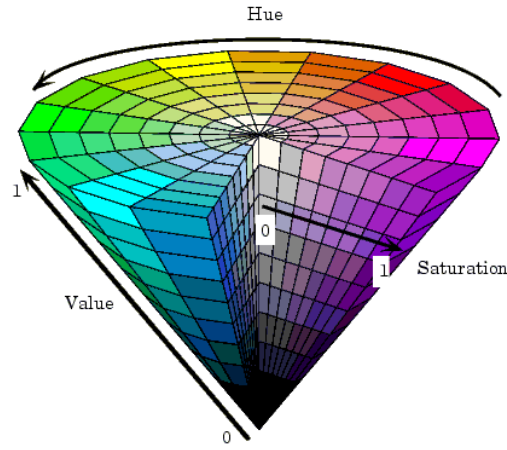
Como visto no artigo *Introduction to color theory* (John W. Shipman, 2012), outra forma de definir uma cor é através do modelo HSV. Esse modelo tem como canais a matiz (Hue), onde o valor 0 é o vermelho e cada valor possui uma cor pura equivalente, $1/6$ do max é o amarelo, $1/3$ verde, $2/3$ azul entre outras. O segundo canal, saturação (Saturation) descreve o quão branca a cor é, sendo o valor 0 a cor branca e o valor 1 a o valor da matiz. O último canal, o valor (Value) se refere ao quão escura a cor é, sendo 0 o valor preto e quanto maior o valor mais clara iluminada a cor. Esse modelo, por possuir informações diferentes da cor em si, pode ser útil para identificar informações como sombras e iluminação.

2.1.4 Vídeo

Como explicado no livro *Digital Video Processing* [20], um vídeo é composto por uma série de imagens, que pode conter ou não informação de áudio. Essas imagens são definidas como frames. Esse vídeo possui uma quantidade finita de frames e uma taxa de reprodução, que define a quantidade de frames que são mostrados em um segundo. É considerado que para um vídeo possuir uma taxa de reprodução satisfatória esse vídeo necessita ter no mínimo uma taxa de 24 frames/segundo.

2.1.5 Formato imagem/frame

Um vídeo ou uma imagem possui uma resolução, que é definida a partir da quantidade de pixels que são utilizados para representar a imagem. A resolução máxima que uma câmera é capaz de representar é comumente chamada de megapixels.



(a) Modelo HSV

Figura 2.4: Representação das cores no modelo HSV. Imagem retirada do artigo "HSV Color Space"(Rohan Library) [12]

Megapixels	Resolução
2 MP	1600 x 1200
3 MP	2048 x 1536
5 MP	2560 x 1920
6 MP	2816 x 2112
8 MP	3264 x 2468
12 MP	4200 x 2800

Tabela 2.1: Resolução referente aos megapixels. Dados retirados do artigo "What is a megapixel?"(Digicam Help) [9]

A proporção de uma imagem é utilizada para definir o formato de uma imagem. Por exemplo, uma imagem com proporção 4 x 3 pode ter qualquer resolução contanto que a proporção sempre seja mantida. Nesse caso ambas resoluções 1600 x 1200 e 2048 x 1536 seguem essa mesma proporção sendo assim imagens que podem ser utilizadas para impressões em papel, que segue a proporção 4 x 3.

2.1.6 Detecção rápida de objetos

No artigo Rapid Object Detection using a Boosted Cascade of Simple Features por Paul Viola e Michael Jones [22] foi definida uma forma utilizando aprendizagem de máquina para detecção rápida de objetos. Para isso, o artigo realizou três contribuições principais.

A primeira foi a criação de uma nova representação de imagem, a imagem integral, onde a imagem pode ser analisada utilizando uma tabela de soma da área. Essa tabela possui o valor em cada pixel como sendo a soma de todos os pixels anteriores verticalmente e horizontalmente, seguindo um sentido da esquerda para a direita e de cima para baixo. Com essa representação é possível utilizar as imagens integrais para verificar o tipo do objeto sendo analisado.

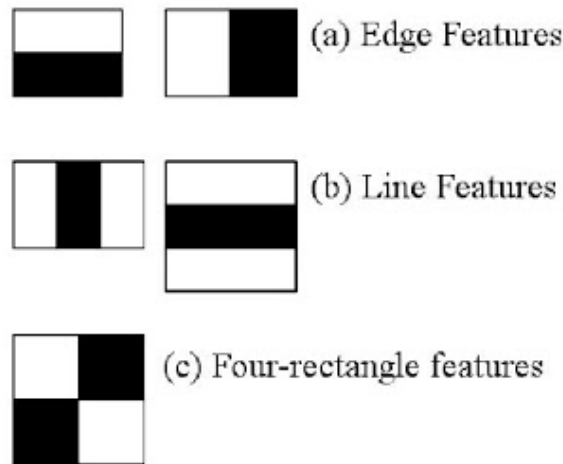


Figura 2.5: Exemplos de imagens integrais. Imagem retirada da documentação do OpenCV[18]

A segunda foi uma metodologia de criação de um classificador, utilizando algumas características selecionadas utilizando o algoritmo AdaBoost. Para isso a metodologia entende, que para analisar um certo objeto o algoritmo deve ser capaz de analisar apenas as características que são consideradas críticas assim diminuindo o tempo de realização da verificação.

E a terceira contribuição foi a utilização de uma metodologia em cascata para focar em regiões, que parecem ser mais interessantes para identificação de um objeto e focando a atenção do algoritmo nessas regiões promissoras agilizando, assim, ainda mais o processo.

2.1.7 Média e Variância em imagem

O valor da média entre os pixels de uma imagem, pode ser utilizado para melhorar contraste enquanto a variância pode ser utilizada para encontrar ruídos, como analisado no artigo "Digital Image Enhancement and Noise Filtering by Use of Local Statistics"(Jong-Sen Lee) [11].

A média pode ser feita utilizando médias locais ou média total da imagem sendo a primeira a que gera informações mais interessantes. Essa média simples,

$$\mu = \frac{1}{n} * \sum_{i=1}^n (x_i) \quad (2.1)$$

pode ser utilizada para analisar o contraste da região e realizar ajustes de contraste.

A variância necessita que a média seja calcula primeiramente pois seu valor significa o valor do desvio quadrático médio da média,

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \quad (2.2)$$

Por essa equação analisar o desvio em uma região, seu valor é útil para encontrar ruídos em imagens, tendo em vista que valores altos de variância significam que existem variações altas na região analisada, assim existindo, com uma probabilidade alta, ruídos na região observada.

2.1.8 iPhone

O iPhone é um dispositivo móvel desenvolvido pela Apple e que ocupa uma grande parte do mercado de celulares global. Os modelos mais novos são o iPhone 7, 7 plus e o SE. Esses modelos de celular utilizam o sistema operacional proprietário iOS que já está em sua versão 10.2. No entanto, uma grande fatia do mercado ainda possui celulares mais antigos como o iPhone 4s. Por tal motivo, e por conta do iPhone 4s já possuir uma câmera de 8 megapixels e um processador dual core de 1GHz[3], esse dispositivo se mostra uma escolha interessante para o desenvolvimento de soluções que buscam qualidade de resultados aliado a um bom desempenho.

2.2 Padrões de fotografia para documentos oficiais

Um padrão de fotografia como o ICAO 9303[1] consiste de características necessárias para que uma fotografia esteja conforme o padrão especificado. Diversos órgãos que utilizam esse padrão foram analisados em busca de encontrar as partes adotadas que são similares. Sendo esses órgãos:

- Padrão da embaixada dos estados unidos [17]
- Padrão criado pela união Europeia [4]
- Padrão para documentos da Universidade de Auckland [16]

Através da análise dos órgãos é possível perceber que algumas características são sempre utilizadas e que, além de algumas características mais simples estarem contidas dentro de outras, existem muitas partes do padrão ICAO 9303[1] que podem ser analisados de forma automatizada através de uma ferramenta, como podem ser observados abaixo.

2.2.1 Posicionamento da face

Todos os padrões analisados definem que a pessoa que está sendo fotografada deve olhar diretamente para a câmera de forma centralizada e contendo a cabeça completa, com cabelo, se a pessoa o tiver.

2.2.2 Iluminação e fundo

Os padrões comparados definem que uma foto deve possuir uma iluminação apropriada, sendo que alguns desses padrões especificam o que é essa iluminação, de tal forma que a imagem tem que estar com coloração correta e que seja fácil visualizar qualquer característica da face do fotografado.

Além desses fatores, os padrões definem que o fundo têm que ser claro e sem ruídos,

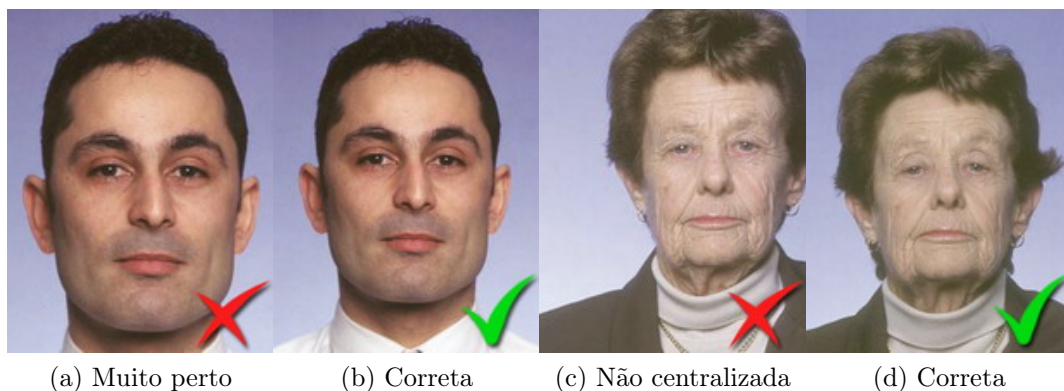


Figura 2.6: Comparação de posicionamento das faces. Imagens retiradas do guia "ICAO Guide for MRTD Photo Guidelines"[10].

sendo que o padrão da embaixada americana define a cor branca como a correta e o padrão europeu a cor cinza claro.

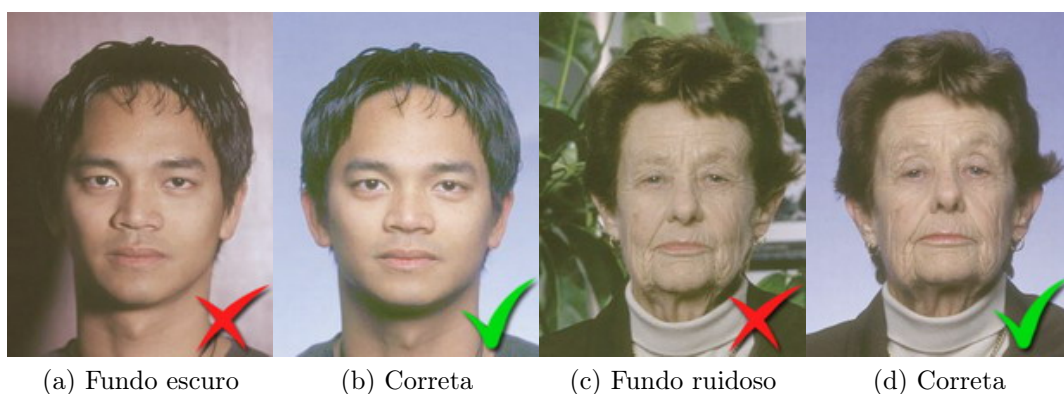


Figura 2.7: Comparação do fundo das imagens. Imagens retiradas do guia "ICAO Guide for MRTD Photo Guidelines"[10].

2.2.3 Formato da imagem

Nesse ponto os órgãos divergem bastante, o Conselho Europeu utiliza fotos 7:9 e a embaixada americana utilizada um padrão 2:2. Um pós processamento da imagem, onde o usuário pode escolher para que tipo de padrão a foto seria utilizada, resolveria esse caso.

2.2.4 Informações adicionais

Como é possível observar na figura 2.8, todos os padrões permitem o uso de óculos contanto que este possua lente transparente e que a armação não atrapalhe na identificação de características importantes da face da pessoa.

Acessórios da cabeça são proibidos caso estes não sejam por razões religiosas, nesse caso o chapéu não pode cobrir nenhuma parte da face, não pode gerar sombras na face e todos os cantos da face têm de estar visíveis.

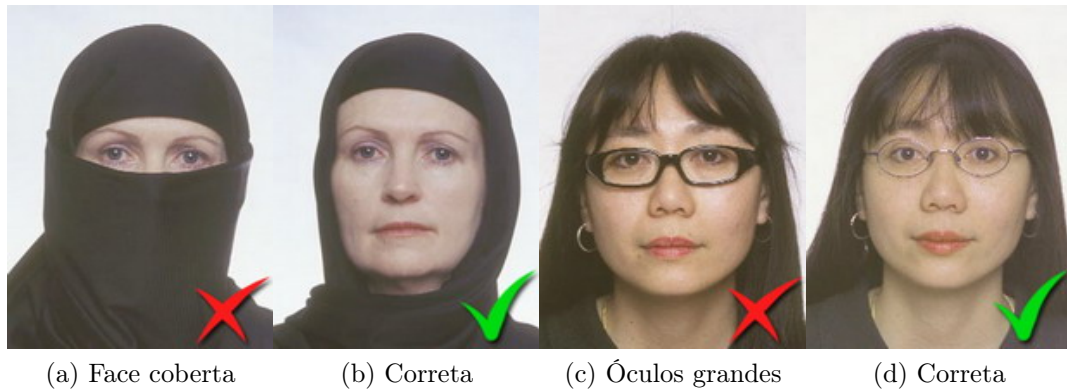


Figura 2.8: Comparação de acessórios nas fotos. Imagens retiradas do guia "ICAO Guide for MRTD Photo Guidelines"[10].

2.3 Trabalhos correlatos

No trabalho Xiufeng Gao, Stan Z. Li, Rong Liu e Peiren Zhang [23] foram definidas algumas formas de avaliar se uma foto foi tirada de maneira que a qualidade da imagem esteja garantida. Essas métricas foram definidas pois, como dito no artigo, a qualidade das imagens de entrada impactam diretamente na qualidade das características, que podem ser analisadas a partir de uma dada imagem. Para que essas formas de avaliação fossem geradas, primeiramente, foram definidos quais são os possíveis erros que a fotografia pode conter, entre eles os mais interessantes para o escopo deste trabalho são os relacionados com o posicionamento da câmera:

- Desvio da pose frontal
- Face muito perto ou muito longe
- Face fora de foco
- Face parcialmente oculta

Com o funcionamento deste trabalho a ideia é que uma pessoa comum tenha a capacidade de tirar uma fotografia padronizada, a fim de evitar que todos os problemas possam ocorrer, para que sistemas como o proposto no artigo [23] possam lidar com menos verificações, agilizando seu processo de avaliação da qualidade das imagens de entrada.

Foi estudado por Oriana Yuridia Gonzalez Castillo [2] como as pessoas classificam as imagens apenas utilizando o olho e como essa classificação pode se dar de forma não padronizada, mesmo quando as pessoas são especializadas no assunto. A pesquisa também aponta a diferença entre uma pessoa especializada e uma pessoa não especializada nessa classificação e como a classificação pode divergir entre o grupo não especializado e o especializado. Um exemplo do artigo é quando para um grupo de imagens as pessoas foram pedidas para classificar as imagens de acordo com 5 critérios:

- A imagem têm boa resolução
- A face é reconhecível

- A cor é neutra
- A imagem está dentro dos critérios necessários
- A imagem fornece um bom desempenho para o software que eu uso

Esses critérios geraram os seguintes resultados:

Pergunta	Usuários com experiência prática	Usuários com experiência técnica
1	25%	11%
2	30%	16%
3	0%	0%
4	35%	46%
5	10%	25%

Tabela 2.2: Percepção de qualidade por usuários experientes

Como é possível notar na tabela 2.2 a análise, mesmo se baseando em critérios específicos, pode divergir consideravelmente, dando assim um uso interessante para o software planejado para o trabalho onde essa classificação se daria de forma automática.

Chee Kiat Ng [15] apresenta um sistema que pode fazer o reconhecimento facial de uma pessoa utilizando um PDA. O PDA é um computador portátil que ficou famoso nos anos 2000 antes do crescimento dos smartphones. O sistema desenvolvido por Chee utiliza de filtros de correlação para identificar a face da pessoa utilizando a câmera enquanto, para encontrar a face, assim como nesse trabalho, Chee optou por utilizar Haar em cascata desenvolvido por Viola-Jones [22]. O motivo dessa escolha é sua agilidade ser suficiente para que ele funcione em dispositivos móveis sem causar grandes atrasos.

Um problema apontado por Chee que foi percebido ao longo desse trabalho é de que para um reconhecimento facial mais preciso e rápido é necessário que a iluminação na face do usuário se dê de forma constante, algo que não costuma acontecer quando um usuário está utilizando o sistema sem ajuda de terceiros.

Capítulo 3

Metodologia

Este capítulo dedica-se a apresentar a metodologia utilizada para aquisição de fotografias padronizadas bem como a forma que essa padronização é classificada.

3.1 Dimensões

Para realizar a captura da imagem é definido primeiramente qual a proporção que será utilizada, como por exemplo 7:9, e então os frames são analisados sempre com a imagem cortada de tal forma que a proporção esteja correta. Dessa forma, é possível aceitar os diferentes modelos de celulares disponíveis, assim suportando diferentes resoluções de imagem, no entanto mantendo sempre a proporção adequada.

3.2 Obtenção dos frames

A estratégia utilizada para capturar uma fotografia padronizada foi utilizar o sistema do celular, câmera frontal ou principal, e realizar a captura do vídeo a 24 frames por segundo e classificar o conteúdo em tempo real. Com esse vídeo foi adotada a estratégia de que a variação para frames próximos seria pequena e que para otimizar o desempenho no celular poderia ser analisada uma característica diferente para cada frame. Caso todas características num intervalo de 10 frames sejam validadas o sistema então bate uma fotografia que seria considerada como válida dentro das métricas estabelecidas.

3.3 Análise dos frames

Para realizar a análise da imagem foi utilizada um sistema onde cada frame possui apenas uma característica. Essa característica então é analisada como um processo e avaliada para que o processamento não interfira na experiência do usuário.

Essa metodologia foi adotada pois testes utilizando uma estratégia em pacote onde cada frame passa por todas validações mostraram que um celular não era capaz de rodar em tempo real sem que tivesse problemas de travamentos e de desempenho.

Como é possível verificar no fluxograma 3.1, é possível seguir essa metodologia pois cada processo/validação depende sempre de apenas uma validação anterior ou nenhuma. Isso só é possível pois podemos considerar que a variação em um intervalo de 10 frames,

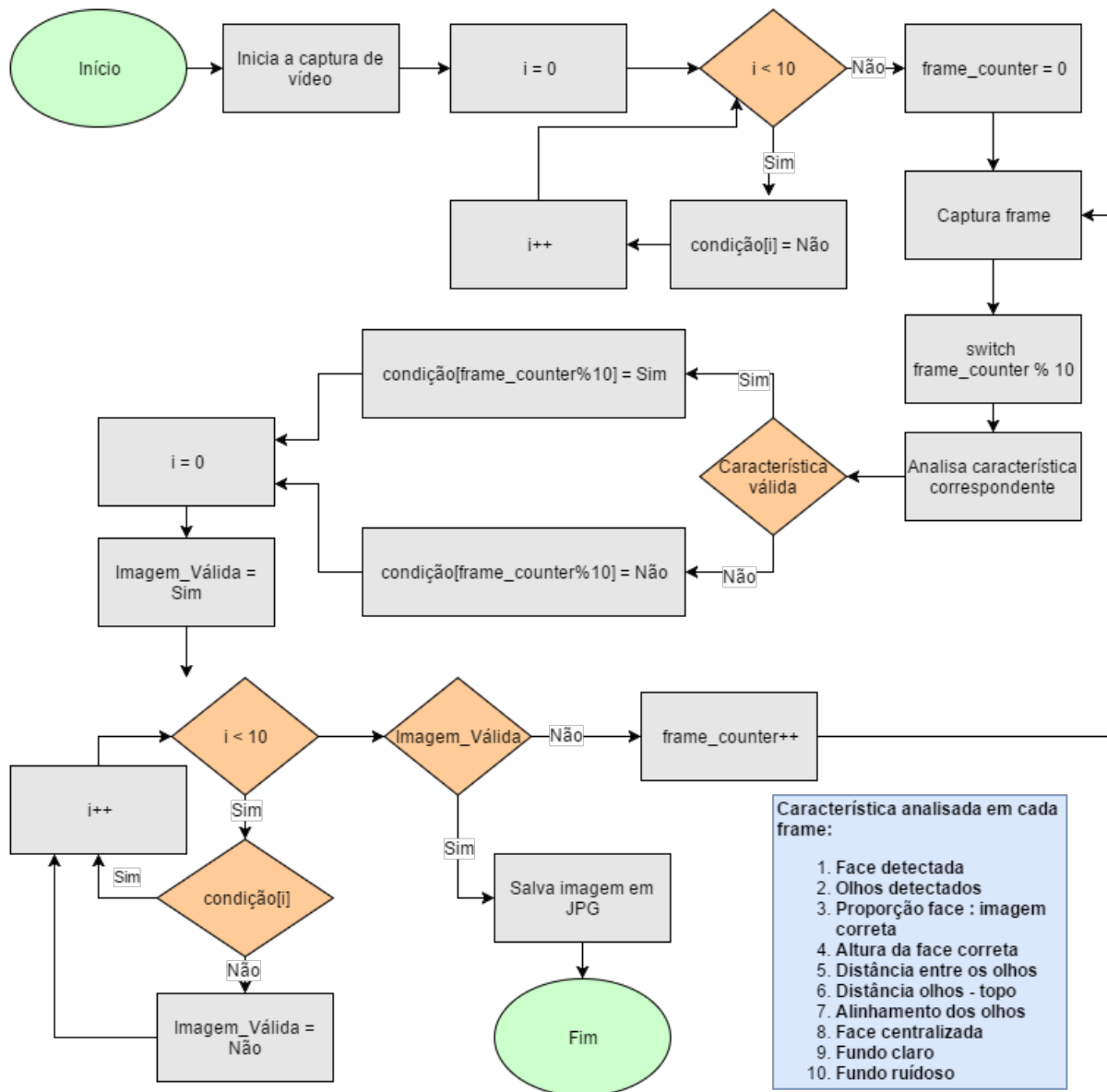


Figura 3.1: Fluxograma proposto do sistema

em uma filmagem de 24 frames/segundo, quando a pessoa está com o objetivo de tirar uma foto para um documento oficial, é muita baixa, pois só pode ocorrer em um intervalo de 0.41s e a pessoa está tentando manter a câmera estável.

3.4 Detecção de faces

Para realizar a detecção de faces foi utilizada a função `detectMultiScale` do OpenCV. Tal classificação é dada como válida caso seja encontrada apenas uma cabeça na imagem. Caso seja encontrada mais de uma cabeça esse frame é invalidado e um alerta é dado ao usuário sobre o problema.

3.5 Detecção de olhos

Para realizar a detecção de olhos foi utilizado a função `detectMultiScale` do OpenCV novamente, no entanto, com um set de características de olhos. A classificação é dada como válida caso apenas dois olhos sejam encontrados. Essa classificação pode ser feita de duas formas:

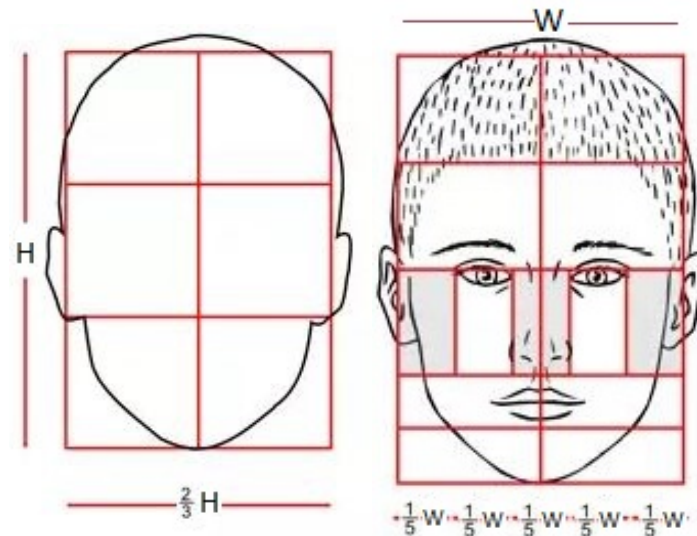
3.5.1 Com face

Caso uma face tenha sido encontrada no passo anterior o algoritmo utiliza a região encontrada e verifica se existem dois olhos dentro da região definida. Caso não encontre, uma mensagem de alerta é dada ao usuário.

3.5.2 Sem face

Caso nenhuma face tenha sido encontrada no passo anterior o algoritmo é aplicado em toda região do frame para tentar encontrar olhos. Caso apenas dois olhos sejam encontrados é possível estimar a cabeça a partir da posição dos olhos levando em conta a premissa de que apenas um usuário está tentando utilizar o sistema proposto.

Para estimar a cabeça utilizando os olhos podemos definir de acordo com o artigo "The proportions of the head and face"[8], que a largura da cabeça é 3 vezes a distância dos olhos somado com suas larguras e que a altura da cabeça é 50% maior do que a largura da cabeça encontrada. O algoritmo 1 proposto demonstra como essas medidas podem ser utilizadas para estimar a face.



(a) Proporções da cabeça, largura : altura
(b) Largura baseada na distância dos olhos

Figura 3.2: Proporções da face. Figura adaptada do artigo "The proportion of the head and face"[8].

Data: Imagem inicial e vector com olhos

Result: Área estimada da face

```

if olho[0].x < olho[1].x then
    olhoEsquerdo = olho[0];
    olhoDireito = olho[1];
else
    olhoDireito = olho[0];
    olhoEsquerdo = olho[1];
end
distanciaOlhos = abs((olhoEsquerdo.x + olhoEsquerdo.width) - olhoDireito.x);
centroFace = centro dos olhos;
larguraFace = 3*distanciaOlhos + olhoDireito.width + olhoEsquerdo.width;
alturaFace = 1.5*larguraFace;
xFace = centroFace.x - larguraFace/2 ;
yFace = (olhoEsquerdo.y + olhoEsquerdo.altura/2) - alturaFace/2 ;
if qualquer borda da face fora da imagem then
    corrigir borda para ficar dentro da imagem
end

```

Algorithm 1: Estimação da face através dos olhos

3.6 Altura da face

Com a face detectada por meio de um dos métodos, é verificada se a altura da face encontrada está correta. Para isso, verificasse se a razão do tamanho da face pelo tamanho da imagem está dentro do limite aceitado entre 50% a 70% o tamanho da fotografia. Esse

limite foi definido através de testes práticos para se encontrar o valor no qual a face se encontra melhor posicionada.

3.7 Distância entre olhos

Com os olhos detectados é avaliado se os olhos estão distantes no máximo metade da largura da fotografia. Tal medida ajuda a saber se a face está a uma distância correta da câmera. Esse teste então avalia se a distância dos olhos está dentro do limiar que é considerado como aceitável para a fotografia. O teste pode ser realizado a partir do algoritmo 2.

Data: Imagem inicial, vector com olhos e máxima distância (maxDist)

Result: Booleano indicando se dentro da faixa

```
centroOlho1 = Point(olho1.x + olho1.largura/2, olho1.y + olho1.altura/2);
```

```
centroOlho2 = Point(olho2.x + olho2.largura/2, olho2.y + olho2.altura/2);
```

```
diff = centroOlho1 - centroOlho2;
```

```
distância = sqrt(diff.x * diff.x + diff.y * diff.y);
```

```
return (distância < maxDist);
```

Algorithm 2: Distância entre os olhos válida

3.8 Distância dos olhos ao topo

Para saber se a face está na posição correta, o Conselho Europeu[4] estabelece que os olhos devem estar a uma distância de 13 a 23 do topo para uma imagem com altura de 45, utilizando uma imagem com proporção 35:45 ou 7:9.

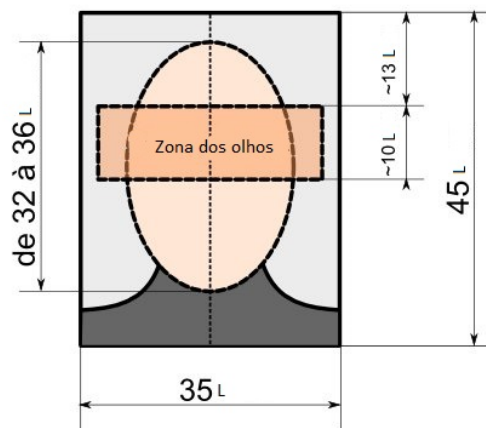


Figura 3.3: Distância dos olhos ao topo. Imagem adaptada de "Photos d'identité [24]."

Portanto, é necessário que o algoritmo verifique se o centro dos olhos se encontra dentro dessa faixa para validar a posição dos olhos. Para isso foi criado o algoritmo 3.

Data: Imagem inicial e vector com olhos

Result: Booleano indicando se dentro da faixa

$centroOlhos = (olhos[0].y + olhos[0].height + olhos[1].y + olhos[1].height)/2;$

$minHeight = 13/45 * imagem.height;$

$maxHeight = 23/45 * imagem.height;$

if $centroOlhos \geq minHeight$ **AND** $centroOlhos \leq maxHeight$ **then**

return true;

else

return false;

end

Algorithm 3: Validação da distância dos olhos ao topo

3.9 Alinhamento dos olhos

Para identificar se a face está alinhada foi utilizada uma estratégia de verificar qual a distância vertical entre os centros dos olhos. Tal distância máxima foi definida através de testes práticos como no máximo a $altura_{do\ olho} / 4$ para evitar que a pessoa utilizando a solução possa inclinar o rosto ou a câmera.

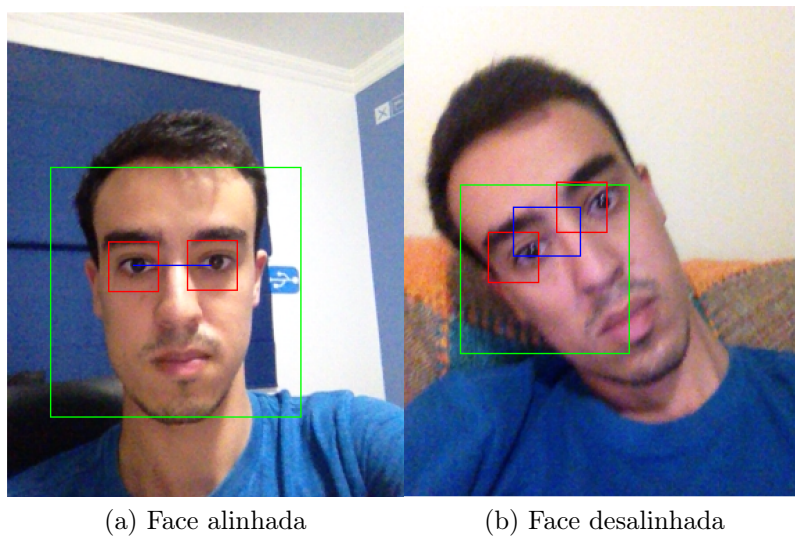


Figura 3.4: O retângulo azul demonstra o alinhamento

3.10 Face centralizada

Tendo em vista que a posição vertical da face já foi verificada na Seção 3.7 é necessário nesse momento verificar se a cabeça está centralizada horizontalmente. Para isso é analisado se o centro da face está dentro de um limiar de distância aceitável do centro da face. Isso pode ser feito utilizando o algoritmo 4.

Data: Imagem inicial, retângulo da face e distância máxima do centro (*maxDist*)

Result: Booleano indicando se dentro da faixa

$\text{faceCenter} = \text{face.x} + \text{face.width}/2;$

$\text{imgCenter} = \text{img.cols}/2;$

$\text{distânciaAbsoluta} = \text{abs}(\text{faceCenter} - \text{imgCenter});$

$\text{distância} = (100 * \text{distânciaAbsoluta})/\text{face.x};$

if *distância* \leq *maxDist* **then**

return true;

else

return false;

end

Algorithm 4: Verificação de face centralizada

3.11 Detecção de fundo

Os padrões definem que para garantir uma boa qualidade de imagem para documentos oficiais o fundo da imagem deve seguir alguns critérios. Portanto, é necessário que o fundo seja analisado de forma separada da imagem para que os requisitos sejam analisados.

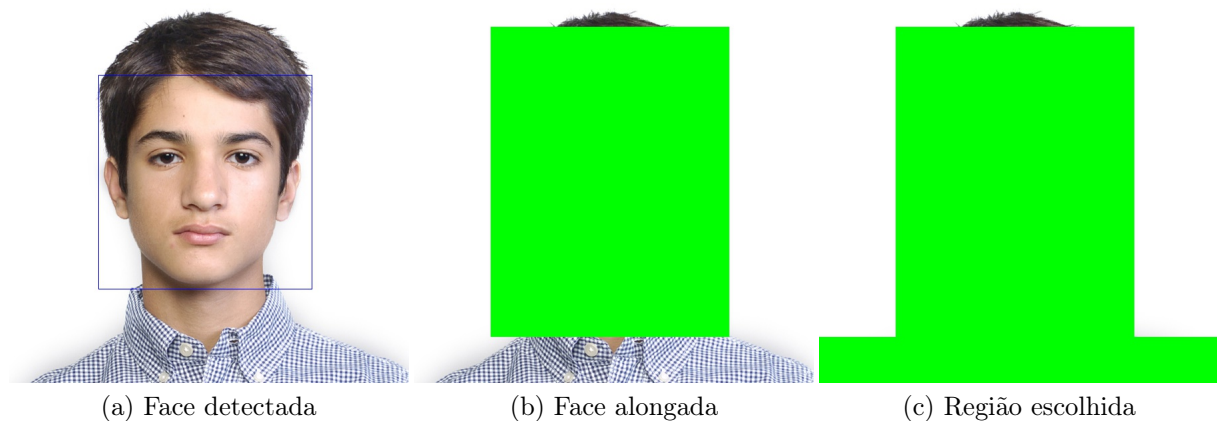


Figura 3.5: Detecção de fundo. Foto exemplo de documento retirada do estúdio de fotos Delio[19]

Para encontrar o fundo é necessário utilizar a região que foi encontrada da cabeça, alongar essa região, para evitar que informações que não são fundo sejam removidas, e depois, remover o pedaço abaixo da cabeça que são os ombros e pescoço da pessoa. A região a ser

analisada pode ser encontrada apartir do algoritmo 5.

Data: Imagem inicial, retângulo da face, deltaAumento

Result: Pontos que identificam o fundo

```
novox = retangulo.x - 0.4 * deltaAumento * retangulo.largura;  
novaLargura = novox + retangulo.largura + 0.8 * deltaAumento *  
    retangulo.largura;  
novoy = retangulo.y - 1.5 * deltaAumento * retangulo.altura;  
novaAltura = novoy + 3 * deltaAumento * retangulo.altura;  
novoRetanguloCabeça = retangulo(novox,novoy,novaLargura,novaAltura);  
fundo = tudo acima de novoy + novaAltura que não esteja dentro do  
    novoRetanguloCabeça;
```

Algorithm 5: Definição do fundo da imagem

3.11.1 Fundo claro

Para verificar se o fundo é claro foi verificado se o fundo encontrado possui em sua maioria cores claras. Para isso a imagem foi primeiramente convertida para o espaço de cor HSV. Com esse espaço de cor, então, foi possível obter, pixel a pixel, o valor de saturação, o valor 'value' e calcular a média dos valores obtidos. Caso esses valores se encontrem dentro dos limiares estabelecidos, significa que a imagem possui um fundo claro válido. Portanto, o algoritmo 6 criado para realizar tal análise.

Data: Imagem inicial, retângulo da face, limiar de saturação máximo(maxSat) e limiar de mínimo de valor(minVal)

Result: Booleano indicando se dentro da faixa

```
valorMáximo = 255;  
hsvImage = hsv(imagem) saturaçãoTotal = 0; valorTotal = 0; totalPixels = 0;  
for i in 0..hsvImage.linhas do  
    for j in 0..hsvImage.colunas do  
        if j > retangulo.x AND j < retangulo.largura AND i > retangulo.y AND i  
            < retangulo.altura then  
            end  
            if i < retangulo.altura then  
                HSVPixel = hsvImage[i][j];  
                saturaçãoTotal += HSVPixel.saturação;  
                valorTotal += HSVPixel.valor;  
                totalPixels++;  
            else  
            end  
        end  
    end  
end  
percentualSaturacao = 100 * (saturaçãoTotal/totalPixels) / valorMáximo;  
percentualValor = 100 * (valorTotal/totalPixels) / valorMáximo;  
return (percentualSaturacao <= maxSat AND percentualValor >= minVal);
```

Algorithm 6: Verificação de fundo claro

3.12 Ruído do fundo

Para garantir que a fotografia feita esteja dentro dos padrões estabelecidos é necessário que a pessoa esteja na frente de um fundo liso, sem muitas variações. Essa variação pode ser analisada utilizando a média dos valores dos pixels e a variância entre esses pixels. Para descobrir a media podemos utilizar o algoritmo 7.

Data: Imagem inicial, retângulo da face, rgbImage

Result: Média encontrada

média = 0;

totalPixels = 0;

for *i* in 0..*rgbImage.linhas* **do**

for *j* in 0..*rgbImage.colunas* **do**

if *j* > *retângulo.x* AND *j* < *retângulo.largura* AND *i* > *retângulo.y* AND *i* < *retângulo.altura* **then**

end

if *i* < *retângulo.altura* **then**

 média += *rgbImage*<float>[*i*][*j*];

 totalPixels++;

else

end

end

end

média = média/totalPixels;

return média;

Algorithm 7: Cálculo da média dos pixels do fundo

Com a média calculada podemos então calcular a variância e analisá-la para verificar se o valor está dentro do limiar desejado como demonstrado no algoritmo 8.

Data: Imagem inicial, retângulo da face, rgbImage, media, varMax, perVarMax

Result: Percetual de variância encontrada

```

altaVariancia = 0;
ruído = 0;
xcentro = imagem.colunas/2;
conta = 0;
blur(imagem.mul(imagem), fotoMedia2, Size(3,3));
blur(imagem, fotoMedia, Size(3,3));
fotoMedia = fotoMedia.mul(fotoMedia);
fotoVariance = fotoMedia2 - fotoMedia;
for i in 0..rgbImage.linhas do
    for j in 0..rgbImage.colunas do
        if j > retângulo.x AND j < retângulo.largura AND i > retângulo.y AND i
            < retângulo.altura then
            end
            if i < retângulo.altura then
                if fotoVariance.at<float>(k, j) > varMax then
                    altaVariancia = altaVariancia + 1;
                    ruído = ruído + (float)abs(xcentro - j) / xcentro;
                    conta++;
                end
            else
            end
        end
    end
end
perVarCalc = 100.0 * ruído / conta;
return (perVarCalc <= perVarMax);

```

Algorithm 8: Verificando a variância dos pixels do fundo

3.13 Salvando a imagem

Após todas as avaliações serem validadas o sistema então apresenta uma tela onde o usuário pode salvar a imagem capturada. Essa imagem é salva no formato JPEG sem compressão para a galeria do celular do usuário e assim pode ser utilizada como quiser.

3.14 Análise dos resultados

Para realizar a análise dos resultados foi escolhida a metodologia de comparar os resultados obtidos pelo software com uma avaliação subjetiva através de uma pesquisa realizada com entrevistados não especialistas. Essa pesquisa foi realizada utilizando 50 imagens onde as pessoas tinham que avaliar de 1 (Completamente fora do padrão) a 5 (Completamente dentro do padrão) a foto tirada com relação a imagem base 3.6.

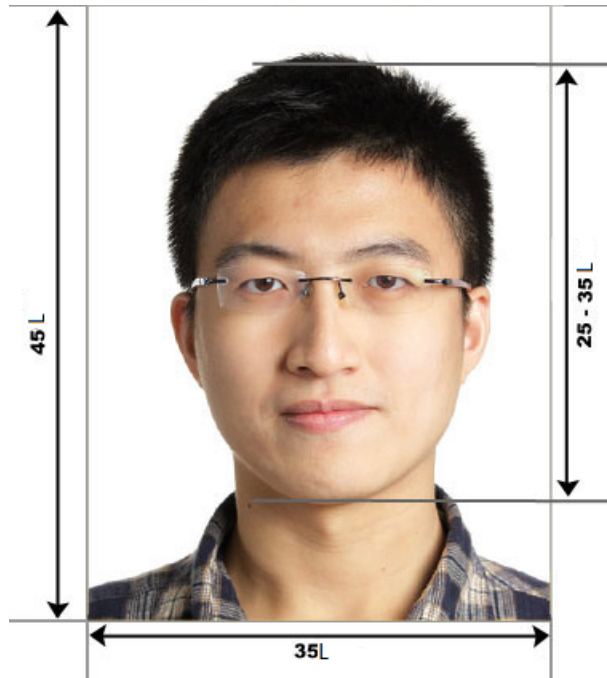


Figura 3.6: Imagem base utilizada na pesquisa. Adaptada de [7].

A obtenção das imagens a ser avaliadas foram feitas utilizando 10 pessoas, sendo 5 imagens de cada pessoa e cada imagem com alguns critérios sendo desabilitados. Os tipos de imagem gerados foram:

- Todos as verificações ativadas
- Alinhamento dos olhos desligado
- Altura da face desligado
- Fundo desligado
- Tudo desligado

Capítulo 4

Resultados

Nesse capítulo serão demonstrados os resultados experimentais de cada um dos sistemas de verificação implementados bem como o resultado da solução funcionando como um todo e avaliada por pessoas não especializadas externas ao trabalho.

4.1 Detecção de faces

Como é possível observar na figura 4.1, a utilização do classificador Haar em cascata é suficiente para detectar a face, no entanto, o quadrado determinado corta alguns pedaços, comumente um pouco do queixo e um pouco da testa.

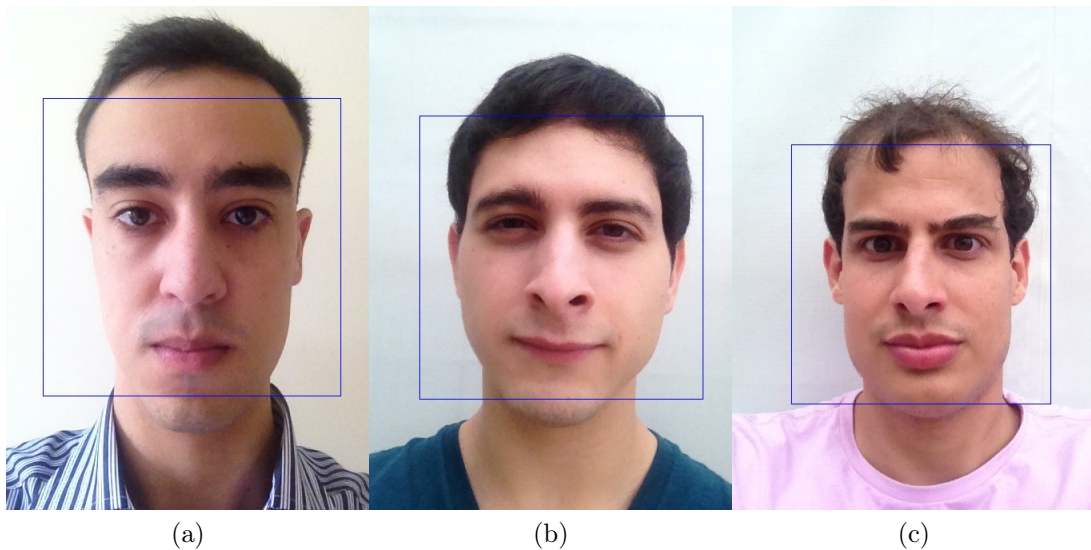


Figura 4.1: Detecção de faces utilizando o classificador Haar em cascata.

4.2 Detecção de olhos

4.2.1 Com face

Caso o sistema já tenha encontrado uma face, são procurados os olhos apenas no retângulo onde a face foi determinada, agilizando assim o processo. Como é possível observar na figura 4.2, os olhos são encontrados de forma correta quando os usuários estão em posição frontal com a câmera.

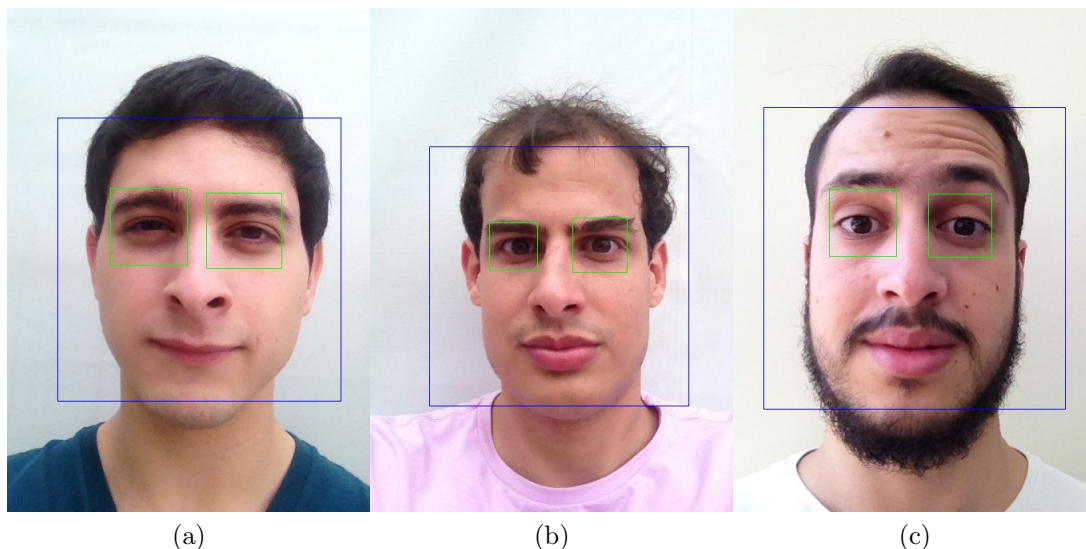


Figura 4.2: Detecção de olhos utilizando classificador Haar em cascata na região da face encontrada.

4.2.2 Sem face

Caso o sistema de detecção de face falhe, ainda é possível que exista uma face no local, portanto, primeiramente são buscados olhos na imagem para então verificar se existe uma possível face com resultados demonstrados na figura 4.3.

Após a detecção dos olhos é possível estimar a cabeça através do algoritmo proposto. Como é possível verificar na figura 4.4, devido ao algoritmo considerar as proporções padrões de uma cabeça, o retângulo estimado da face acaba sendo mais preciso do que o encontrado através da detecção direta da face, quando essa se encontra de maneira frontal.

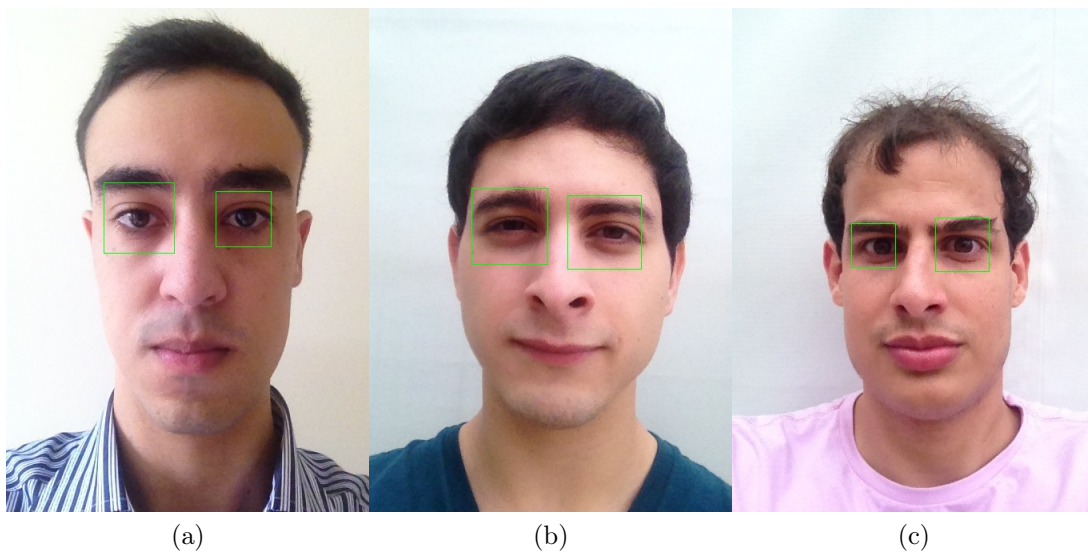


Figura 4.3: Detecção olhos utilizando classificador Haar em cascata.

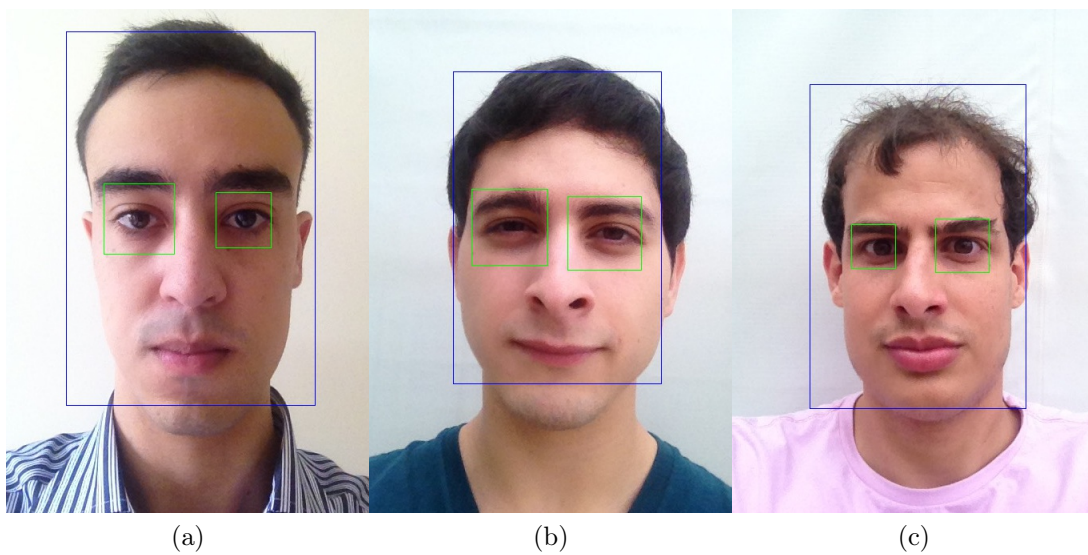


Figura 4.4: Estimação da face através dos olhos encontrados.

4.3 Altura da face

A razão definida como correta para a face é que sua altura seja entre 50% á 70% da altura total da fotografia. Valor definido analisando os padrões estudados. Portanto, é possível utilizar as imagens anteriormente encontradas para realizar a comparação de quais estariam dentro do padrão.

Imagem	Razão	Dentro do padrão
4.4a	74.28	Não
4.4b	62.06	Sim
4.4c	64.44	Sim
4.1a	59.21	Sim
4.1b	56.35	Sim
4.1c	51.59	Sim

Tabela 4.1: Razão da face em relação a imagem.

Como é possível observar na tabela 4.1, com a aproximação através dos olhos, seria encontrada uma altura mais precisa no caso da imagem 4.4a que resultaria em uma foto gerada pelo sistema mais adequada ao padrão desejado.

4.4 Distância entre olhos

Foi definido no sistema que a distância entre os centros dos olhos têm que ser no máximo a largura da imagem dividida por 2 para que essa seja considerada uma imagem válida dentro dos padrões estudados. Tendo em vista que o sistema verifica antes a razão do tamanho da face com o tamanho da imagem está dentro de um limiar, essa verificação é útil apenas para detectar possíveis inconsistências.

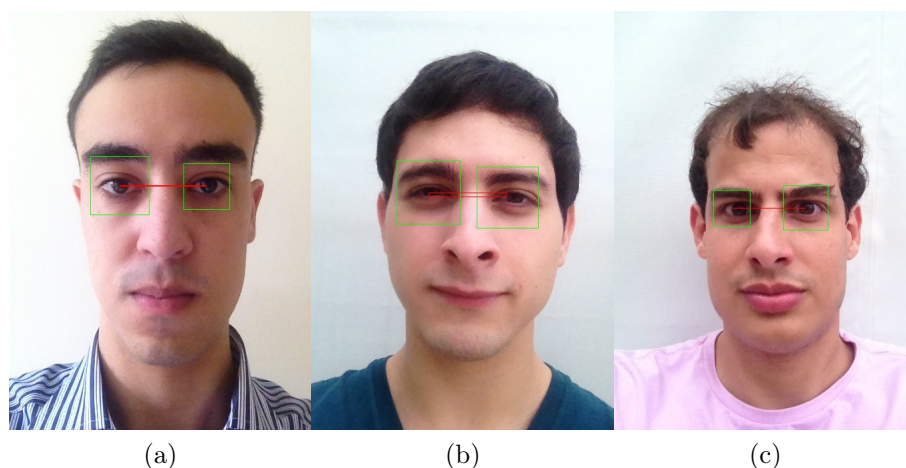


Figura 4.5: Distância entre o centro dos olhos.

Temos assim na figura 4.5 exemplos de imagens com distâncias entre os olhos válidas e uma demonstração de como a distância entre os olhos é determinada.

4.5 Distância dos olhos ao topo

Foi definido a partir do padrão adotado que o centro dos olhos deveria estar a uma distância no máximo $13.0/45.0 * altura da imagem$ e no mínimo $23.0/45.0 * altura da imagem$. Portanto, temos na tabela 4.2 os seguintes resultados para as imagens analisadas, tendo todas uma altura de 630 pixels, sendo 182 a distância mínima e 322 a distância máxima.

Imagem	Distância	Dentro do padrão
4.5a	263	Sim
4.5b	276	Sim
4.5c	296	Sim

Tabela 4.2: Distância dos olhos ao topo.

4.6 Alinhamento dos olhos

Para verificar o alinhamento dos olhos foi utilizado o retângulo criado na verificação da distância entre o centro dos olhos. Foi analisada então se a altura desse retângulo é menor que a $1/4$ da altura do olho com maior altura. Caso isso seja verdade, o alinhamento é considerado como correto. Temos assim como

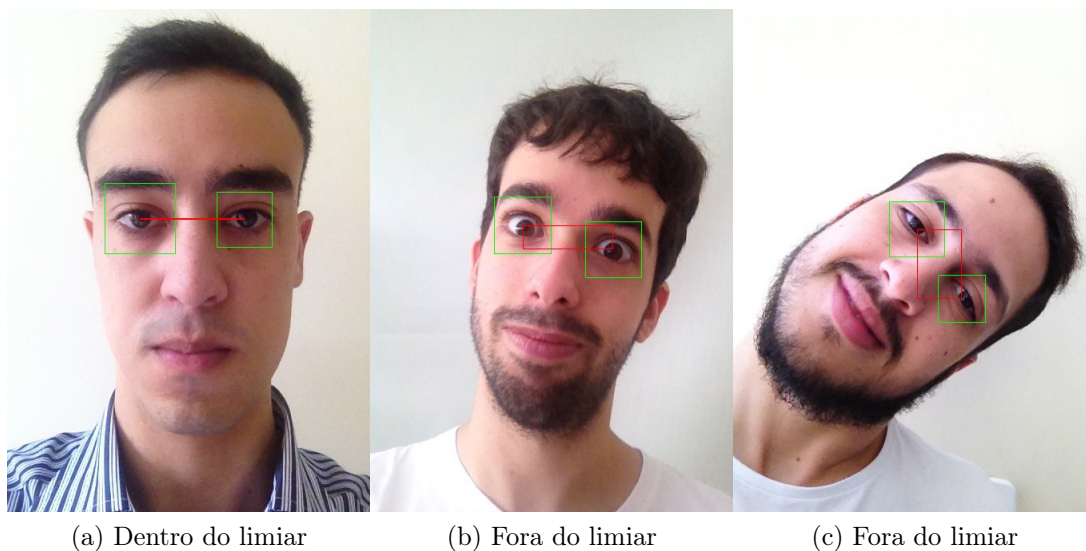


Figura 4.6: Comparação de alinhamento dos olhos.

Como é possível observar, as figuras 4.6b e 4.6c são consideradas como fora do padrão evitando assim que fotografias que não poderiam ser utilizadas em documentos oficiais sejam aceitas.

4.7 Face centralizada

Para garantir que a face está centralizada foi verificado se o centro da face se encontra a uma distância no máximo de 10% do centro da imagem. Esse valor foi definido pois após realizar os testes práticos e as necessidades de uso do sistema uma margem maior de erro para o usuário resultaria em um sistema mais simples de utilizar. A medida de 10% é estabelecida levando em conta que o centro é o ponto com 0% de deslocamento e a borda da imagem o ponto com 100%.

Figura	Deslocamento	Dentro do padrão
4.4a	1%	Sim
4.4b	3%	Sim
4.4c	2%	Sim

Tabela 4.3: Tabela demonstrando a centralização da faces.

Analisando a tabela 4.3 podemos observar que as imagens tiradas utilizando o sistema apesar de terem um certo desalinhamento do centro ainda foram aceitas por estarem dentro do limiar estabelecido.

4.8 Detecção de fundo

A detecção de fundo, como estabelecido na metodologia, é realizada através de uma projeção do tamanho total da face e de uma remoção dos espaços a baixo da cabeça para não considerar a roupa da pessoa como fundo. Portanto, como pode ser observado nas figuras 4.7 e 4.8 a forma com que a cabeça foi identificada resulta em estimações do fundo diferentes.

4.8.1 Face encontrada

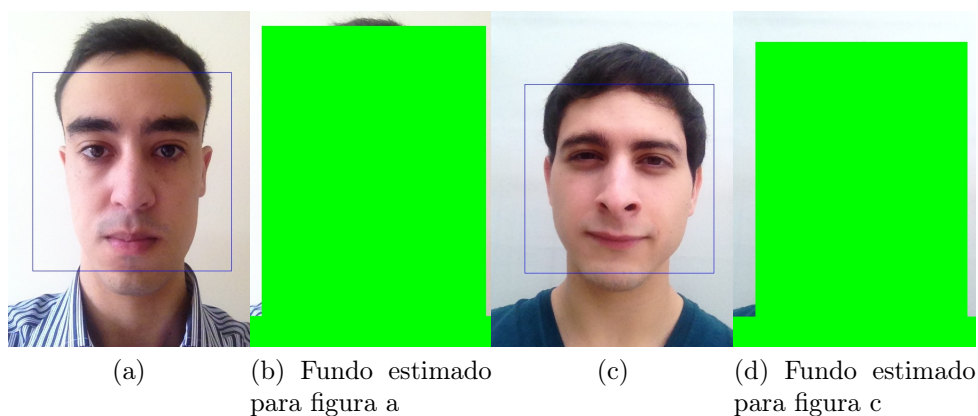


Figura 4.7: Demonstração de como o fundo é encontrado com a face encontrada.

4.8.2 Face estimada pelos olhos

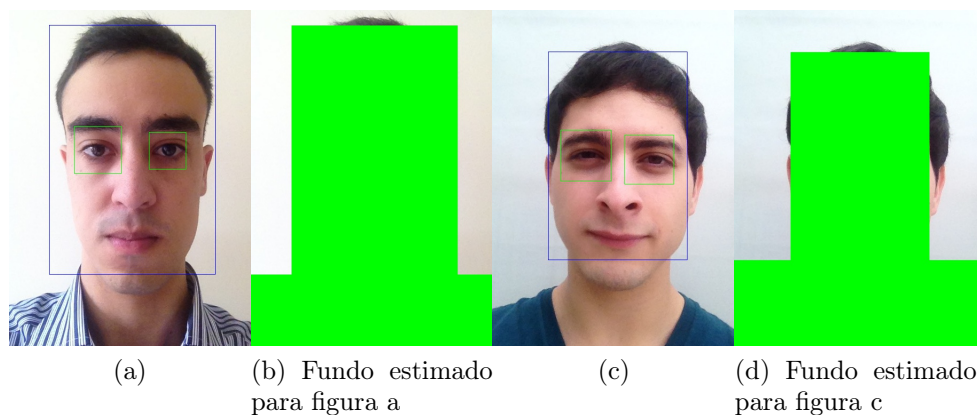


Figura 4.8: Demonstração de como o fundo é encontrado com a face estimada.

Podemos ainda perceber que o fundo foi melhor estimado com a face sendo descoberta através dos olhos comparando as figuras 4.7 e 4.8. Isso pode trazer benefícios para a análise do fundo realizada no sistema, criando assim uma menor possibilidade de que fundos não desejados sejam aceitos.

4.9 Fundo claro

Para verificar se o fundo de uma imagem é claro foi utilizado o espaço de cores HSV. Com esse espaço de cores foi definido que para o fundo ser claro ele necessita ter um valor médio de saturação a baixo de 10% e um valor de cor acima de 90%. Para descobrir esses valores foi utilizado uma ferramenta de edição de cor e analisado qual era o limiar máximo que deveria ser aceito no sistema.

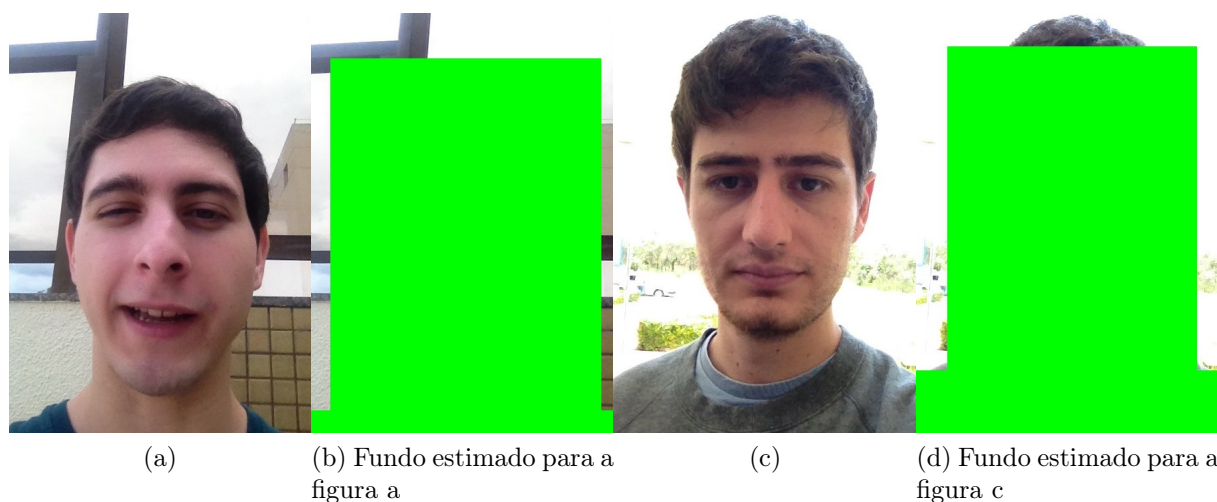


Figura 4.9: Exemplos de imagens tiradas com o fundo colorido.

Figura	Saturação	Valor	Dentro do padrão
4.8a	4.41%	92.96	Sim
4.8c	6.63%	94.16	Sim
4.9a	7.14%	85.05	Não
4.9c	2.85%	94.08	Sim

Tabela 4.4: Verificação de claridade do fundo.

Como é possível observar na tabela 4.4 mesmo que em algumas imagens o fundo possua cores e ruídos o que está sendo analisado é a claridade e portanto a figura 4.9a foi a única classificada como com fundo escuro, estando assim o sistema funcionando conforme o esperado.

4.10 Ruído do fundo

Para verificar se uma imagem possui ruído no fundo foi utilizada a medição da variância do fundo da imagem, caso essa variância fosse de valor superior a 10% então o fundo seria ruidoso demais para ser aceito pelo sistema.

Figura	Variância	Dentro do padrão
4.8a	0%	Sim
4.8c	0%	Sim
4.9a	6.64%	Sim
4.9c	5.34%	Sim

Tabela 4.5: Verificação de ruído do fundo.

4.11 Sistema completo

O sistema funciona, como pode ser observado na figura 4.10, através do processamento em cada frame de uma característica. Caso alguma característica de filtragem esteja fora do padrão desejado uma mensagem é mostrada para o usuário para que ele possa corrigir o problema e continuar tentando tirar uma fotografia dentro do padrão. Caso todos os filtros retornem um resultado positivo, o sistema então, automaticamente, captura uma foto que é considerada como uma foto válida dentro dos padrões desejados.

Para analisar o funcionamento do sistema foi criada uma base de 49 imagens, a partir de 10 pessoas, que foi avaliada por 541 pessoas não especialistas. Para realizar essa avaliação as pessoas foram dadas uma imagem exemplo, figura: 4.11, uma descrição sobre o trabalho e uma pergunta fixa para cada imagem. A pesquisa foi dividida em 5 pesquisas para que cada pessoa necessitasse avaliar apenas 10 imagens.

Descrição da enquête:

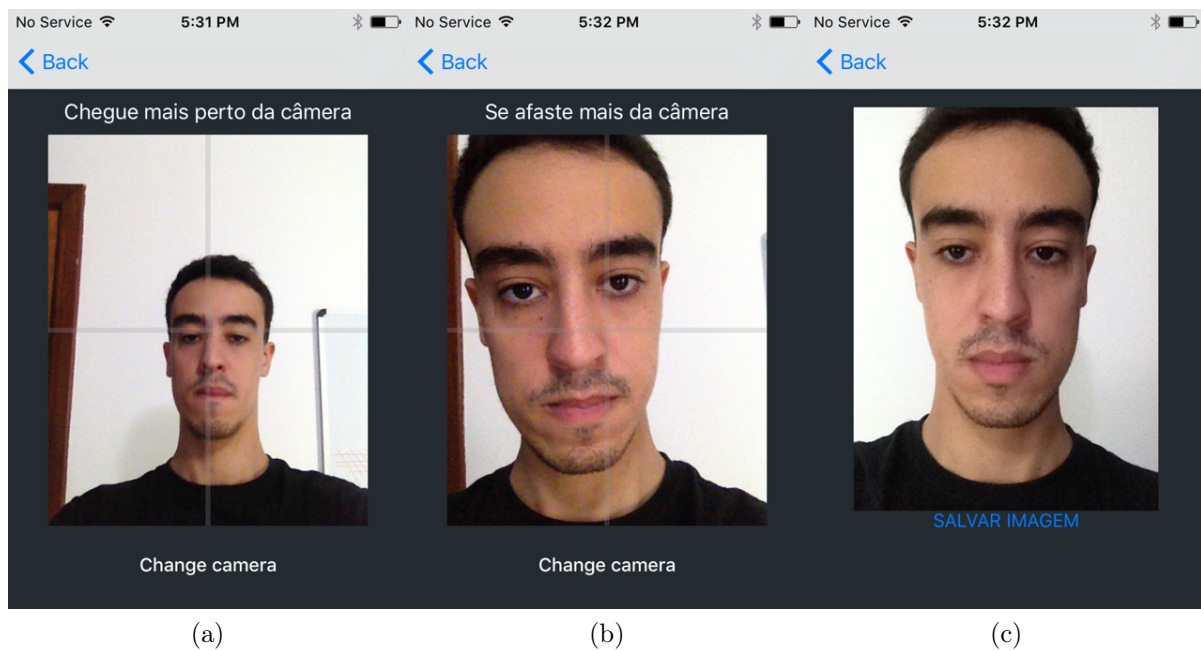


Figura 4.10: Imagens do sistema funcionando.

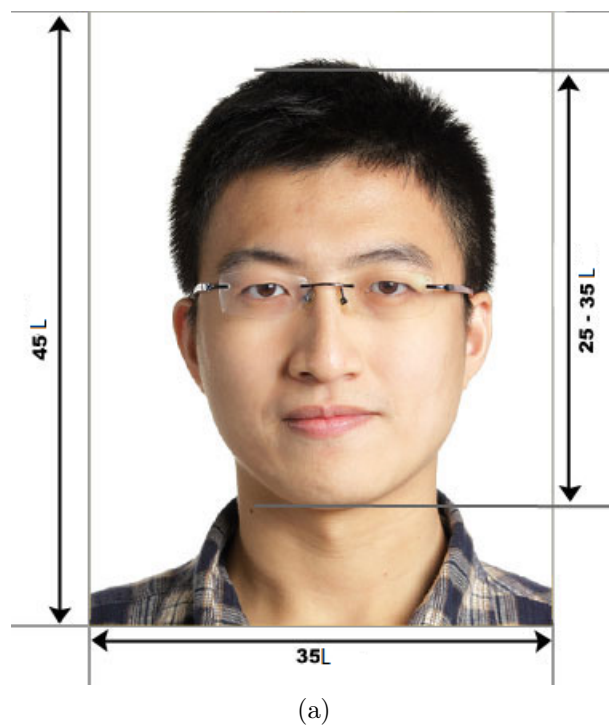


Figura 4.11: Imagem de referência. Adaptada de [7]

O objetivo do trabalho é desenvolver um sistema que ajude as pessoas a tirar fotos para documentos oficiais sem a necessidade de ajuda de outra pessoa.

Pergunta: De 1 a 5, o quanto essa foto está dentro do padrão desejado? (1 Completamente fora) (5 Completamente dentro)

Com os diferentes tipos de fotos que foram tiradas para cada usuário com o sistema ativo, completamente ou parcialmente, podemos então verificar o ganho que o sistema traz para os usuários.

4.11.1 Todas as verificações desativadas x Todas as verificações ativas

As figuras 4.12b e 4.13b são alguns exemplos das fotos que foram tiradas para comparar a qualidade do sistema. Para isso é necessário que exista um ganho real entre a qualidade das fotos tiradas com o sistema funcionando e sem o sistema funcionando garantindo assim uma motivação para seu uso.

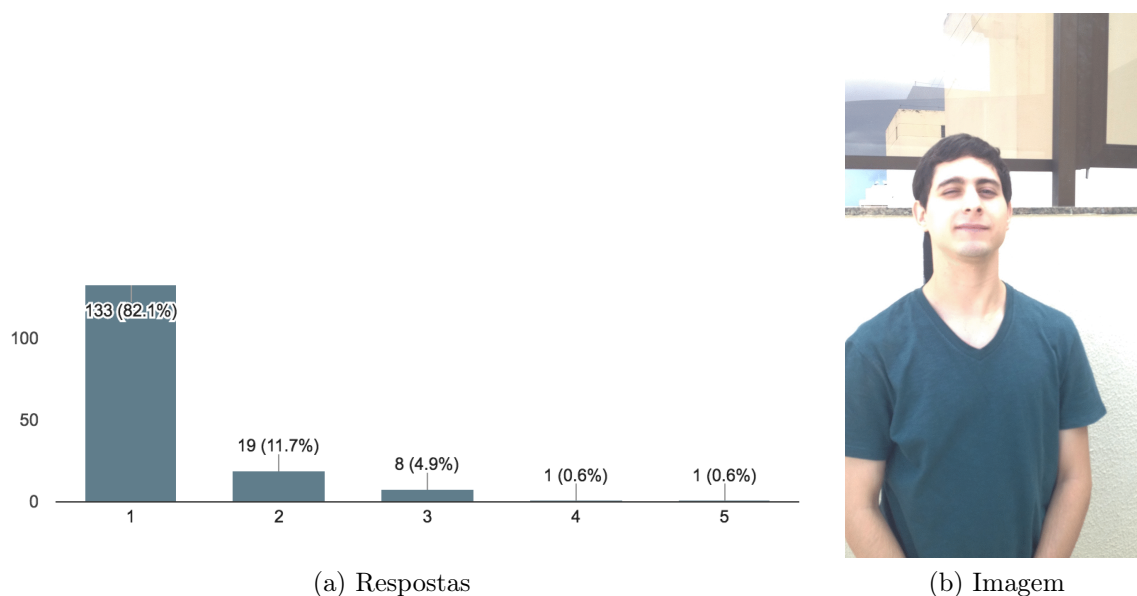
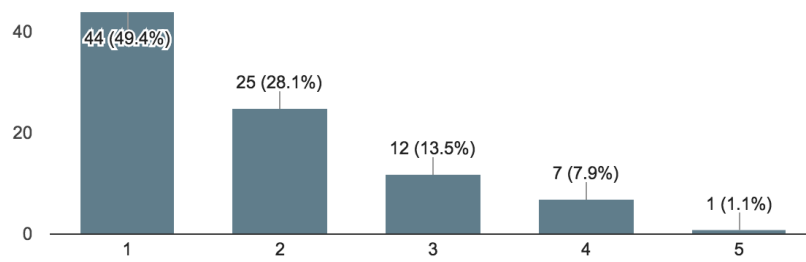


Figura 4.12: Imagem 1 com tudo desligado.

Como é possível observar, com o sistema funcionando temos um percentual de rejeição pelos entrevistados não especializados da figura 4.12b de 93.8% e da figura 4.13b de 80% se analisarmos os valores abaixo de 3 como rejeitados.

Temos então as figuras 4.14b e 4.15b como exemplos de fotografias tiradas com o sistema todo funcionando. Com isso, os usuários do sistema foram limitados a seguir as dicas dadas e as fotografias só foram tiradas quando elas passavam nos testes realizados.

Com isso temos uma classificação de 70.2% para a figura 4.14b como uma foto aceita e de 61.8% para a figura 4.15b. Podemos perceber então que ao utilizar o sistema uma melhoria significativa no sistema que pode ser detectada mesmo por pessoas não especializadas quando perguntadas.

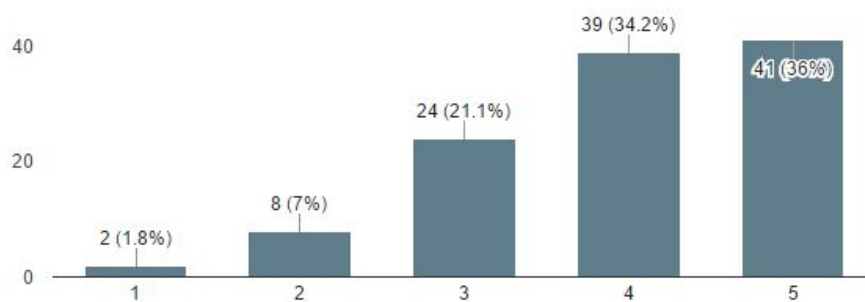


(a) Respostas

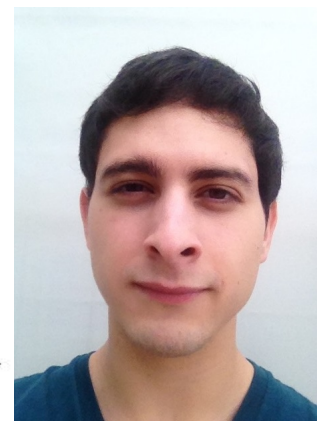


(b) Imagem

Figura 4.13: Imagem 2 com tudo desligado.

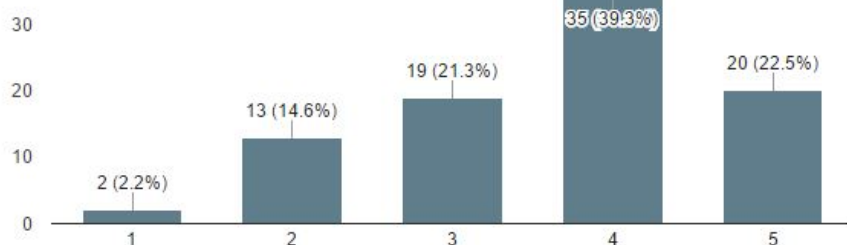


(a) Respostas



(b) Imagem

Figura 4.14: Imagem 1 com tudo funcionando.



(a) Respostas



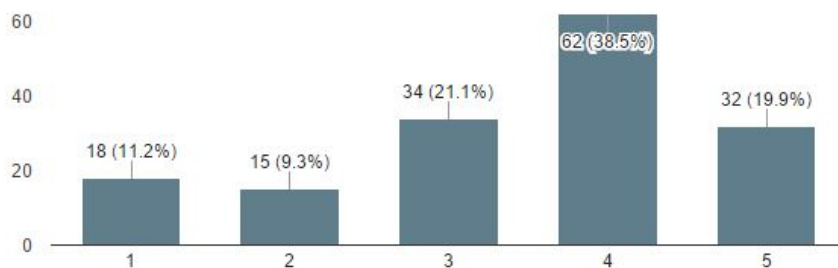
(b) Imagem

Figura 4.15: Imagem 2 com tudo funcionando.

4.11.2 Fotos com sistema de verificação do fundo desligado

As fotos com fundo desligado foram realizadas para verificar se o sistema funcionaria bem caso o fundo não fosse um critério importante e para entender como seria as respostas para imagens com fundo diferente do padrão.

Observa-se então na figura 4.16b uma aceitação de resultados de 58.4% e uma reprovação de 20.5%. Já para a figura 4.17b temos uma aceitação de 52.5% e uma reprovação de 22.3%. Comparando esses resultados com os resultados anteriores podemos observar que os usuários classificaram as imagens como boas ou muito boas assim mas com menor taxa do que com o sistema todo ativado. Isso aconteceu pois os entrevistados não são especializados e não foram capazes de perceber os problemas gerados pelos fundos.



(a) Respostas



(b) Imagem

Figura 4.16: Imagem 1 com fundo desligado.

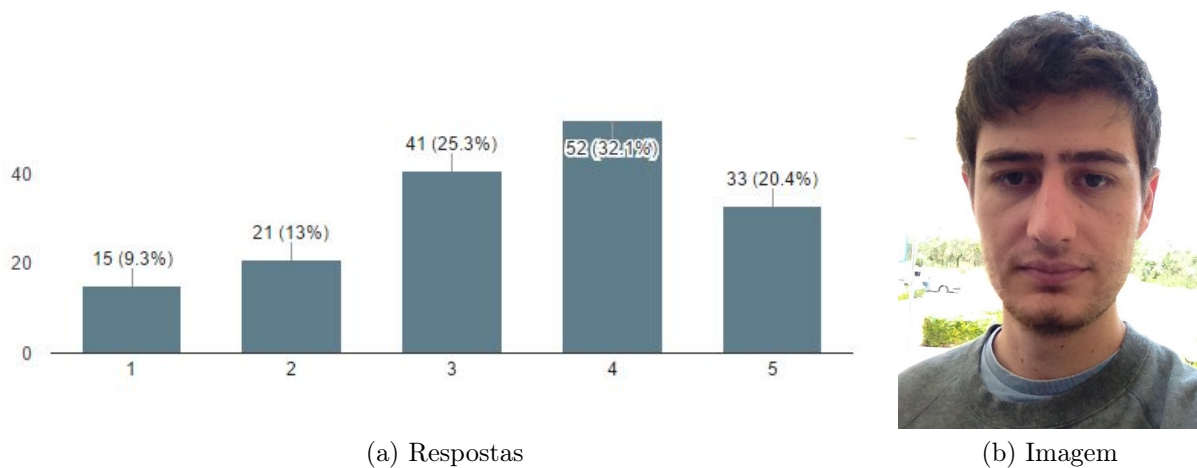


Figura 4.17: Imagem 2 com fundo desligado.

4.11.3 Fotos com sistema de verificação da altura da face desligado

As fotos com verificação do tamanho da face foram analisadas com intuito de descobrir o quanto esse sistema é importante para a solução proposta. Para realizar a análise foi pedido para os usuários tirar fotos com a câmera longe da face e verificar se as pessoas conseguiam perceber esse problema.

Observa-se na figura 4.18b que quando o braço da pessoa aparece na foto os entrevistados conseguiram facilmente identificar, com taxa de reprovação de 87%, que a foto estava fora do padrão, no entanto, quando apenas a cabeça estava pequena na foto, na figura 4.19b, a taxa diminuiu para 56.2%. Comparando com os resultados com o sistema completamente desabilitado e com o sistema habilitado percebemos que o ganho proporcionado pela análise do tamanho da face é relevante e o sistema é necessário.

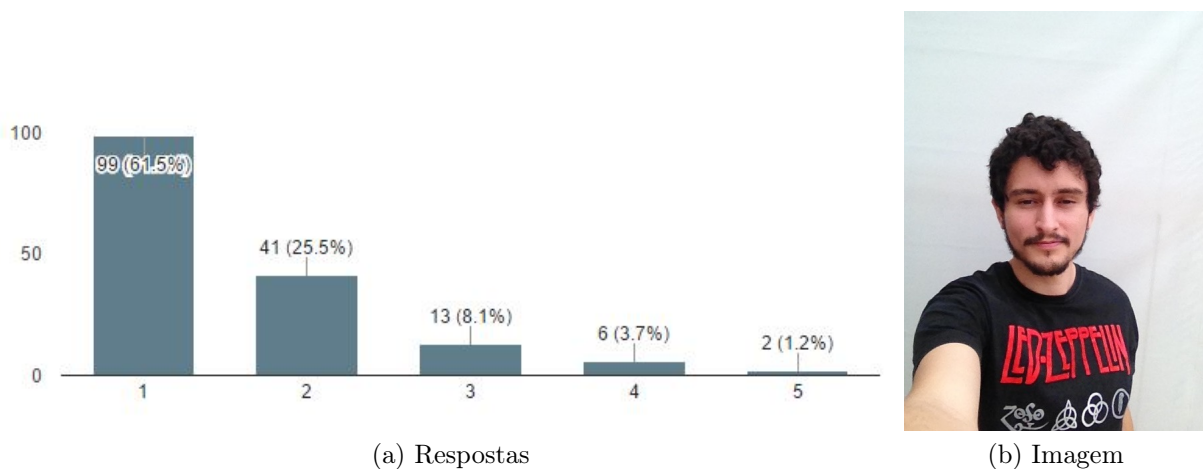
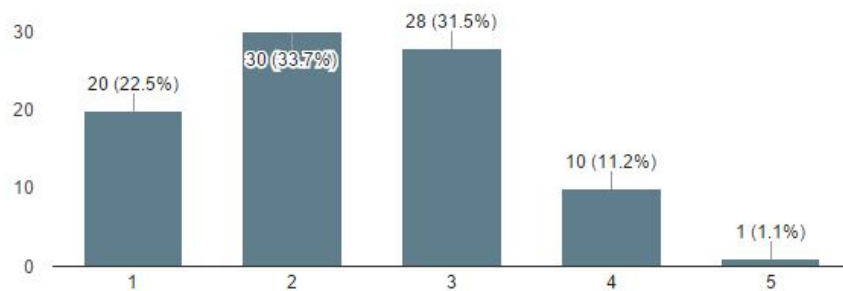


Figura 4.18: Imagem 1 com tamanho da face desligado.



(a) Respostas



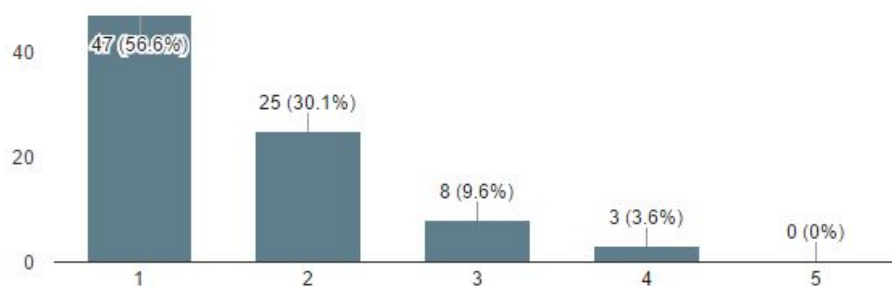
(b) Imagem

Figura 4.19: Imagem 2 com tamanho da face desligado.

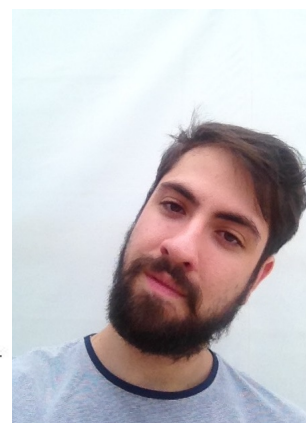
4.11.4 Fotos com sistema de verificação o alinhamento dos olhos desligado

Outro sistema que é importante ser verificado sua importância é o de alinhamento dos olhos. Para isso esse sistema foi desligado e fotografias dos usuários foram realizadas.

Na figura 4.20b verificamos uma taxa de reprovação de 86.7%. Já para a figura 4.21b temos uma taxa de reprovação mais baixa, 48.2% e um percentual central de 34.9%. Podemos então que o ganho do sistema de alinhamento é alto e necessário quando comparado com o sistema todo funcionando e quando comparado com o sistema desabilitado.

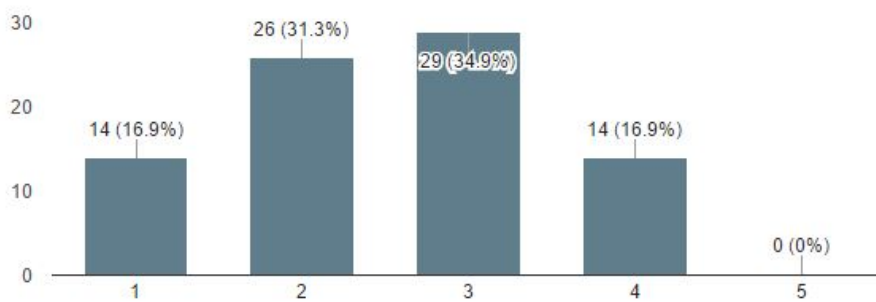


(a) Respostas



(b) Imagem

Figura 4.20: Imagem 1 com alinhamento desligado.



(a) Respostas



(b) Imagem

Figura 4.21: Imagem 2 com alinhamento desligado.

Capítulo 5

Conclusão

O trabalho consistiu de identificar padrões utilizados por documentos oficiais, escolher um padrão que é amplamente utilizado, estudar e desenvolver formas de identificar se uma foto está de dentro do padrão adotado e desenvolver um sistema onde o celular é capaz de tirar fotos dentro do padrão sem muito conhecimento do usuário. Os resultados do trabalho indicam que a solução proposta foi alcançada mas que é possível melhorar o sistema em algumas áreas cruciais para que ele possa ser adotado por organizações que necessitam de fotos padronizadas. Entre essas melhorias temos uma verificação de iluminação e sombras da imagem para garantir uma imagem com menor chances de problemas e uma melhoria na interface utilizada pelo usuário.

Em comparação com os sistemas que existem atualmente para os dispositivos Android e iOS esse sistema se mostrou mais robusto e melhor preparado para o uso das pessoas quando essas têm como fim gerar uma foto para um documento oficial. Para documentos que não necessitam de uma padronização de fotografias tão rígidas como inscrições para provas e concursos o sistema já é suficiente e melhoraria muito a qualidade das fotos recebidas pelas empresas responsáveis pela realização de tais avaliações.

O sistema ainda possui algumas fragilidades na identificação de olhos que poderiam ser melhoradas para que a fotografia padronizada fosse tirada de forma mais rápida e o reconhecimento de expressões faciais ainda não é feito, o que pode ser algo necessário para o futuro.

Para dar continuidade ao desenvolvimento desse trabalho seria necessário então, estudar métodos para analisar expressões faciais, melhorar o sistema de reconhecimento de olhos e procurar formas de avaliar a iluminação e sombras que possam danificar a qualidade da fotografia padronizada.

Referências

- [1] specifications for machine-readable passports, visas and official travel documents. Standard, ICAO, 2015. 8
- [2] Oriana Yuridia Gonzalez Castillo. Survey about facial image quality. *Fraunhofer Institute for Computer Graphics Research*. 10
- [3] Tudo Celular. Apple iphone 4s - ficha técnica. <http://www.tudocelular.com/Apple/fichas-tecnicas/n2196/Apple-iPhone-4S.html>. 8
- [4] Conselho da União Europeia. Regulamento 2252/2004 do conselho. 1, 8, 16
- [5] Tomi T. Ahonen e Alan Moore. *Communities Dominate Brands*. futuretext, 2005. 1
- [6] Rafael C. Gonzalez e Richard E. Woods. *Digital Image Processing*, volume 4. Prentice-Hall,Inc., Upper Saddle River,New Jersey 07458, 2 edition, 2002. 3, 4
- [7] Kodak express. <http://www.kodakexpresscamden.com/passport-singapore.html>. vii, 22, 31
- [8] Arty Factory. The proportions of the head and face. <http://www.artyfactory.com/portraits/pencil-portraits/proportions-of-a-head.html>, 2016. vii, 14, 15
- [9] Digicam help. What is a megapixel? <http://www.digicamhelp.com/camera-features/camera-parts/megapixels/>. ix, 6
- [10] ICAO. Icao guide for mrted photo guidelines. http://www.icao.int/Security/mrted/Downloads/TechnicalReports/Annex_A-Photograph_Guidelines.pdf. vii, 9, 10
- [11] Jong-Sen Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2(2), 1980. 7
- [12] Rohan Library. Hsv color space. <http://www-rohan.sdsu.edu/doc/matlab/toolbox/images/color11.html>. vii, 6
- [13] Microsoft. Rgb color spaces. [https://msdn.microsoft.com/en-us/library/windows/desktop/dd372185\(v=vs.85\).aspx](https://msdn.microsoft.com/en-us/library/windows/desktop/dd372185(v=vs.85).aspx). vii, 5
- [14] Andro Nepal. Passport id photo maker studio. <https://play.google.com/store/apps/details?id=np.com.njs.autophotos&hl=en>, 2014–2016. 1

- [15] Chee Kiat Ng. <https://www.ece.cmu.edu/research/publications/2005/cmu-ece-2005-007.pdf>. *IEEE Xplore*. 11
- [16] The University of Auckland. Id card photo requirements. <https://www.auckland.ac.nz/en/about/admission-and-enrolment/ae-before-you-arrive/cs-id-cards/cs-id-card-photo-requirements.html>, 2016. 8
- [17] U.S. Department of State Bureau of Consular Affairs. Photo requirements. <https://travel.state.gov/content/passports/en/passports/photos/photos.html>, 2016. 1, 8
- [18] OpenCV. Face detection using haar cascades. http://docs.opencv.org/trunk/d7/d8b/tutorial_py_face_detection.html. vii, 7
- [19] Delio Photo Studio. <http://deliophotostudio.com/miami-passport-photos/miami-passport-photos.html>. vii, 18
- [20] A. Murat Tekalp. *Digital Video Processing*. Prentice-Hall, Inc., Upper Saddle River, New Jersey 07458, 2 edition, 1995. 5
- [21] Tinkertanker. Id photoprint: Us passport photos, delivered. <https://itunes.apple.com/us/app/id-photoprint-us-passport/id558788006?mt=8>, 2015–2016. 1
- [22] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–511–I–518 vol.1, 2001. 6, 11
- [23] Rong Liu Xiufeng Gao¹, Stan Z. Li and Peiren Zhang. Standardization of face image sample quality. *ICB 2007, Seoul, Korea, August 27-29, 2007. Proceedings*, 4642, 2007. 10
- [24] Gilles ZIMMER. Photos d’identité. <http://www.photogz83.com/photos-didentit>. vii, 16